A Retrospective Evaluation of a Community Based Applied Behaviour Analysis Service for Children and Youth with Autism Spectrum Disorder

Evangelo Boutsis, M.ADS.

Centre for Applied Disabilities Studies

Submitted in partial fulfillment of
the requirements for the degree of

Masters of Arts

Faculty of Social Sciences, Brock University
St. Catharines, Ontario

Abstract

Limited research exists on the use of applied behaviour analysis to treat problem behaviour of children with Autism Spectrum Disorder (ASD) in real world clinical settings. A retrospective file-review study was conducted to examine the clinical processes and behavioural outcomes for children and youth with ASD in a community-based ABA service (ABAS) for treating problem behaviour. Intervention plans included the function of the behaviour for 91% of clients. The majority of plans (91%) included prevention, skill development, and intervention strategies. Outcome measurement focused on either functional replacement skills (52%) or problem behaviour (48%). Moderate decreases (57%) in problem behaviour, and substantial increases (84%) in replacement behaviours were observed for percentage change outcomes. Large effect sizes (Cohen's d=>0.8) were observed for decreases in problem behaviour (d= 1.23) and as increases in replacement skills (d=-6.80). Positive and meaningful outcomes were achieved for the vast majority of clients in this sample participating in ABAS.

Acknowledgements

I would first like to thank my thesis supervisor Dr. Rosemary Condillac from the Centre for Applied Disabilities Studies at Brock University. Without Dr. Condillac's unwavering support, vision and encouragement I would not be successfully completing a thesis and especially one that I am proud of about. I am truly appreciative and grateful of the opportunity to be your student and learn from you!

I would like to thank you Ateeq Awan, Lauren Martell, Puneet Sharma and Meagan Scott, without your passionate input and participation, this thesis would not have successfully been completed.

I would also like to thank the Centre for Behaviour Health Sciences and Mackenzie Health and more specifically, Jim Hughes and John Clarke for their accommodation, encouragement and ongoing support of this project.

I would also like to thank Dr. Maurice Feldman, Dr. Jan Frijters and Dr. Kirsten Wirth for their valuable input and feedback on this project.

Finally, I would like to express my profound gratitude to my wife Rosemary and children Nicholas, Alexander and Gianfranco. Thank you for your ongoing support, patience and encouragement throughout the process of completing this thesis. I am not completing this thesis without sacrifices, love and support for you! This accomplishment is not possible with you!
To my children:

"All our dreams can come true, if we have the courage to pursue them."

-Walt Disney

Evangelo Boutsis

Table of Contents

List of Tables

List of Figures

A Retrospective Evaluation of a Community-based Applied Behaviour Analysis Service for Children and Youth with Autism Spectrum Disorder

In Canada, from 2003 to 2010, the prevalence of children diagnosed with an Autism Spectrum Disorder (ASD) has steadily increased (National Epidemiologic Database for the Study of Autism in Canada (NEDSAC, 2012). The Center for Disease control (CDC, 2015) estimates that 1 in 68 children in the U.S. have been identified with an ASD. Applied Behaviour Analysis (ABA) is the treatment of choice for this population, with substantial evidence to support its use (The National Autism Centre, 2015). To meet the needs of this increasing population, the Ontario Ministry of Children and Youth Services (MCYS) introduced funding for intensive behaviour intervention (IBI) in 2000 and has steadily increased those funds over the past 16 years.

In 2011, MCYS announced and funded additional ABA-based services across the Province for children and youth diagnosed with ASD (Office of the Auditor General of Ontario; Autism Services and Supports Report, 2012), given the effectiveness of the intensive ABA programs in Ontario (e.g. Perry et al., 2008). Requirements outlined by MCYS for these ABA-based services included the following: (a) services are to be provided between 2-6 months dependent on child and youth needs; (b) emphasis on parent/caregiver involvement in identifying the goals of their children; (c) Parents/caregivers are to receive information, practise and apply techniques with their children or youth; and (d) services are to be provided in a variety of settings (familiar and novel) such as home, school and other community settings. (http://www.children.gov.on.ca/htdocs/English/topics/specialneeds/autism/guidelines/guidelines-2011). Goals that are targeted in the ABA-based services can include (1) increasing adaptive skills such as communication, activities of daily living, and social skills; or (2) reducing problem

behaviour. More specifically, when services are focused on reducing problem behaviours, caregivers supporting the child are trained and provided with strategies to reduce any problem behaviours and/or increase functional replacement behaviours. Subsequently, the objective in the ABA-based services is to build community capacity by enabling caregivers to support their child now and in the future. The significant investment in ABA-based services included investment in program evaluation. Common in program evaluations is the utilization of standardized clinical measures (Baker-Ericzén, Stahmer, & Burns, 2007; Kovshoff, Hastings & Remington, 2011) and other qualitative measures typical in research (Dyer, Martino & Parvenski, 2006; Sharp, Burrell, & Jaquess, 2014) and are used across mental health services in the province. According to Burns (2014), traditional methods used in program evaluations may not capture the individual features of interventions that contribute to treatment outcomes. Data utilized by behaviour analysts to determine effects of intervention in applied settings provide "the richest" and "most clinically meaningful" information about the effect of interventions, however the use of single subject data is often overlooked in the evaluation of programs (Burns, 2014). Given the heterogeneity of the ASD population, single subject designs provide individual impact from intervention that would not be captured in randomly assigned research (Koegel , Koegel, Ashbaugh, & Bradshaw, 2014) and by standardized tools (Burns, 2014). To meet the need of capturing information from single subject designs used by community clinicians, Condillac (2009a) developed a program evaluation system including a tracking tool called the Behaviour Analysis Treatment Taxonomy (BATT; Condillac 2009d) to synthesize key information pertaining to assessment and treatment approaches and outcomes for individual cases.

Given the limited published research on the clinical application of ABA real world settings, the purpose of this study was to retrospectively evaluate one of the ABA-based services

(referred to as ABAS) by examining the clinical processes used by behaviour consultants and the corresponding behavioural outcomes for the children and youth with ASD who participated in the service. In addition, the retrospective evaluation will only include closed participant cases where reducing a problem behaviour was the reason for referral for services. As a result, this study will expand the literature by providing further evidence of the field-effectiveness of ABA for treating problem behaviour in children and youth with ASD.

Literature Review

The following literature review will begin with a brief overview and impact of problem behaviours often exhibited by children and youth with ASD. Next, an overview of common intervention needs, service activities and processes that are utilized to intervene with these target behaviours will be reviewed. More specifically, this will include an overview of applied behaviour analysis (ABA), common target behaviours, types of assessments, treatment components and service delivery models used in the treatment of problem behaviours exhibited by children with ASD. Finally, an overview will be provided of program evaluations of community based services that are based on a mediator model for children with ASD. This review of the literature will lend support to the rationale for this study, which examined the clinical processes and outcomes in an MCYS funded ABA Service (ABAS) for children and youth with ASD.

Problem Behaviours and ASD

Children diagnosed with ASD have an increased likelihood of exhibiting problem behaviours at some point in their lifetime (Horner et al., 2002; Kanne & Mazurek, 2011; Lecavalier, 2006). These problem behaviours include aggression to others, self-injurious behaviours, destruction to property, temper tantrums (Hattier et al., 2011; Lecavalier, 2006;

Hartley et al., 2008, Horner et al., 2002) and unusual vocalizations (Dixon et al., 2011). Further, the severity of problem behaviour often exhibited by children and youth with ASD can range from mild to severe (Kanne & Mazurek, 2011). The presence of problem behaviours may impact the child's ability to learn (Hartley et al., 2008; Horner et al., 2002); caregiver's well-being (McIntyre et al., 2002; Eisenhower, 2005; Hadgetts et al., 2013), family finances and routines (Hadgetts et al 2013), parent burnout (Hadgetts et al., 2013), staff burnout (Hastings & Brown, 2002), decisions regarding out of home placements (McIntyre et al., 2002), and social and community opportunities (Horner et al., 2002). In addition, Hadgetts et al. (2013) found families living with individuals with ASD who exhibit aggression report feelings of isolation, exhaustion, and are concerned about their safety.

Unfortunately, without effective intervention there is an increased likelihood that problem behaviour will continue to worsen (Horner et al., 2002). Further, interventions based on the principles of behavior analysis have been demonstrated to be effective in the assessment and treatment of problem behaviours exhibited by children and adults with developmental disabilities including autism (Carr & Durand, 1985; Feldman, Tough, Condillac, Hunt, & Griffiths, 2002; Horner, Carr, Strain, Todd, & Reed, 2002; Horner, Day, & Day, 1997; Kennedy et al., 2000; Lang et al., 2010). Similarly, the National Autism Center (2015) identified behavioural interventions (e.g. applied behaviour analysis, behavioural psychology, positive behaviour supports) as the largest category of established interventions for children, adolescents and young adults diagnosed with ASD under the age of 22.

Applied Behaviour Analysis

According to Baer, Wolf & Risley (1968), ABA is the application of principles of learning and behavior in the improvement of socially significant behaviors. More specifically, an

essential dimension of ABA is, "*Applied"* which is the identification of an observable behaviour that is socially significant and important to the individual and to their family (Baer et al., 1968). According to Baer et al. (1968) other integral dimensions of ABA include: (a) *Behavioural*-the objective measurement of the behaviour (b) *Analytical*-use of single subject design to determine functional relationship of intervention and the target behaviour (c) *Technological*-thorough description of the procedures in the behavioural interventions (d) *Conceptually systematic*-based on basic behaviour principals based on the literature (e) *Effectiveness*-determined through the evaluation of the data (f) *generalizable*- improvements in socially significant behaviours occur across other behaviours, people and settings. Furthermore, these dimensions are recommended to be included in ABA based literature and in behavioural intervention plans (Baer et al., 1968).

Identifying and Measuring Target Behaviours

As previously described by Baer, Wolf & Risley (1968), identifying a target behaviour for behavioural intervention that is socially significant is integral in applied behaviour analysis. Target behaviours are identified for intervention and determined to be socially significant by the individual and his or her family with the support of the behaviour analyst (Carr et al., 1999; Feldman et al., 2002). Furthermore, once a target behaviour has been identified for intervention, objective measurement of this behaviour is integral to the development of a behavioural intervention plan. Measurement serves several purposes for behaviour analysts including determining the need for intervention, assessment of behaviour, the evaluation of the effectiveness of interventions, informs decisions about program modifications, and in the determination of mastery of goals (Baer et al., 1968; Lerman, Dittlinger, Fentress, & Lanagan, 2011).

Moreover, determining the dimension of the target behaviour requiring intervention is integral in the selection of an appropriate measurement method to utilize (Kahng et al., 2011). More specifically, dimensions of a target behaviour that may require intervention include the frequency (e.g. how often), duration (e.g. the time to complete tasks), latency (e.g. the time to initiate a task after being instructed) and intensity (e.g. the force of a behaviour) (Kahng et al., 2011). In addition, methods that can be utilized to measure these dimension of the target behaviour include: (a) frequency: continuous recording of the number of responses that occur per unit of time (b) response duration: recording the amount of time in which a target behaviour occurs (Kelly, 1977) (c) response latency: recording of the elapsed time between the onset of a stimulus and initiation of a subsequent response (d) intensity: the recording of magnitude, force, effort of a target behaviour on a predetermined scale (Kahng et al., 2011). In addition, percentage recording can be utilized for all the aforementioned dimensions. More specifically, a ratio is formed by dividing the same dimensional quantities such as count (count/count), and time (latency/latency); (duration/duration); (Cooper, Heron & Heward, 2007). Prior to intervention, behaviour analysts are required to complete relevant assessments and measurement, with consideration of client needs, environmental parameters and other contextual variables (Behaviour Analyst Certification Board: BACB, 2016). Further, once the target behaviour is identified as problematic, a functional behaviour assessment (FBA) must be completed (Horner et al., 2002; BACB, 2016).

Functional Behaviour Assessment

The purpose of an FBA is to identify the function(s) of problem behaviour through the identification of contingencies (e.g. consequences, discriminative stimuli, and establishing operations) that maintain problem behavior (Hanley et al., 2003; Horner & Carr, 1997). Problem

behaviour can have one or more of the following functions: (a) to gain attention from others, (b) to escape from demands, (c) to access to preferred items or activities, (d) automatic reinforcement (Hanley, Iwata & McCord, 2003; Iwata et al., 1994). However, the FBA may yield more than one function for a problem behaviour (Carr et al., 1999; Hanley et al., 2003).

Once completed, the FBA informs the development of a behavioural intervention plan. Determining the function(s) for the problem behaviours is integral to the development of an effective behavioural intervention (Sturmey, 1994). Compared to non-function based interventions, function based interventions are associated with greater improvements to reducing problem behaviour results (Carr et al., 1999; Ingram & Lewis-Palmer, Sugai, 2005). An FBA can include different methodologies, including indirect assessments, descriptive or direct assessments, and experimental assessments (Alter et al., 2008, Carr et al., 1999), which will be reviewed in the following sections.

Indirect assessments. Indirect methods involve the collection of information from parents, teachers, and children through the use of questionnaires, rating scales and record reviews (Alter et al., 2008). Specific indirect assessments include, structured interviews such as the Functional Assessment Interview (FAI; O'Neill et al., 1997), questionnaires such as (a) Questions About Behavior Function (QABF; Vollmar & Matson 1996); (b) Motivational Assessment Scale (MAS; Durand and Crimmins, 1988); (c) Functional Analysis Screening Tool (FAST; Iwata & Deleon, 1996). The benefits of utilizing indirect assessments include the time efficiency of implementation and the potential identification of other important information (Roscoe et al., 2015). Similarly, Hall (2005) noted the benefits of questionnaires as they are both time and cost efficient, however, it was noted that the psychometric properties of these tools are sometimes poor. More specifically, information obtained through the use of indirect methods

may be compromised due dependency on the recollections of caregivers and lack of evaluated psychometric soundness of the measure (Iwata, DeLeon, & Roscoe, 2013; Kelley, LaRue, Roane & Gadaire, 2011).

Direct assessments. Direct or descriptive analyses include the direct observation of the behaviour under naturally occurring conditions that precede and follow the behaviour (English & Anderson, 2006). Descriptive analyses have been utilized in the behavioural sciences for decades (Thompson & Borrero, 2011). Although, outcomes of direct methods does not provide verification of a functional relation between events, in combination or alone these methods can be utilized to develop a hypothesis of the function of behaviour (Thompson & Borrero, 2011). The following will highlight several direct methods utilized in a FBA. The first type of descriptive method is the ABC (antecedent-behaviour-consequence) continuous recording (Bijou, Peterson, & Ault, 1968). This method involves the documentation of the antecedents, behaviours, and consequences in natural settings (Alter et al., 2008). In addition, Freeman, Anderson, and Scotti (2000) developed a structured descriptive analysis where antecedent conditions in the natural environment are arranged to increase the likelihood that problem behaviour events will occur. Finally, a third type of direct assessment is a scatterplot analysis (Touchette, McDonald, & Langer, 1985). A scatterplot analysis can be utilized to document the occurrence of problem behaviour according to the time of day and activities. In contrast with indirect methods, descriptive analysis provides correlational data concerning the relations between environmental variables and behaviour (Oliver et al., 2015). Similar to indirect methods, ABC checklists, scatterplots and other direct methods require very little training, and can be implemented with relatively at a low cost (Thompson & Borrero, 2011). However, similar to

indirect methods, a limitation to direct methods is the inability to confirm a functional relationship or function of the problem behaviour.

Experimental assessment. Experimental analyses or functional analogues are experimental designs involving the systematic manipulation of specific conditions to determine the functional or causal relationship between environmental conditions and a target behaviour (Horner, 1994). Two types of analogues are reviewed in this section, functional analysis and structural analysis. Functional analysis involves the systematic manipulation of both antecedent and consequent conditions that can be contrived in the natural environment to determine the function of the problem behaviour (Iwata et al., 1994). In comparison, structural analysis is the assessment and manipulation of antecedent conditions or discriminative stimuli that set the stage for the target behaviour (Stichter, Sasso & Jolivette, 2004). In comparison to structural analysis, which has received less attention in research (Hagan-Burke, Gilmour, Gerow, & Crowder, 2015; Stichter et al., 2004), there is an abundance of research utilizing functional analysis as part of FBA (Hanley et al., 2003; Hanley, Iwata & Thompson, 2001, Iwata et al., 1994). Further, functional analysis has been demonstrated to be highly effective in identifying variables that maintain problem behaviour (Iwata et al., 1994) and given its rigorous procedure it has been described as the gold standard in which other FBA's are compared to (Hanley et al., 2003). However, according to Gersham, Quinn, & Restori (1999), functional analysis limitations include the amount of time and expertise required to implement this procedure.

Currently, there are no regulations or best practices that specify the particular method to be used when conducting a FBA (Oliver et al., 2015). However, several authors have reported caution and poor results if using indirect methods by themselves to determine function (Alter, Conroy, Mancil & Haydon, 2008; Carr et al., 1999). Similarly, Thompson & Iwata (2007) stated

that given the low correspondence with results from functional analysis, results from descriptive assessments should be viewed with caution.

The utilization of functional analyses by behaviour analysts in practice is inconsistent with the ample research supporting the inclusion of a functional analysis in an FBA. Roscoe, Phillips, Kelly, Farber & Dube (2015) surveyed 205 behaviour analysts in practice and found that few reported using functional analysis. The respondents reported that descriptive assessments were the most common assessment method utilized and often in conjunction with indirect assessments. Similarly, Oliver, Pratt & Normand (2015) surveyed 724 behaviour analysts, and reported that the use of indirect and descriptive assessment were more frequently utilized compared to a functional analysis. In addition, the majority of respondents reported "never" or "almost never" using a functional analysis to determine the function for problem behaviour. In addition, reported barriers to the implementation of functional analysis by behaviour analysts included: were not trained (Ellingson, Miltenberger, & Long, 1999), perceived that functional analysis was the hardest to use (Ellingson et al., 1991) and potentially costly in regards to personnel and time (Carr, 1994).. Similarly, as described by Thompson & Borrero (2011) in practice, treatment is often guided by the results of descriptive analysis alone or in conjunction with indirect methods, as opposed to a functional analysis. The aforementioned review of methods utilized by behaviour analysts are inconsistent with the importance placed on the inclusion of functional analysis in FBA's in the literature.

Targets of Intervention and ASD

As described above identifying socially significant behaviour is integral to the development of behavioural interventions (Baer et al., 1968). Behavioural interventions can be used to increase behaviours (e.g. functional skills) and or to decrease behaviours (e.g. problem

behaviour) that are socially significant to an individual and his or her family (Carr et al., 1999; Feldman et al., 2002, National Autism Center, 2015). The National Standards Project Phase 2 (The National Autism Centre, 2015) in their review of 453 articles revealed the common types of target behaviours targeted in behaviour interventions for children and youth with ASD under the age of 22 from 2007 to 2012. The results indicated that the common targeted behaviours to increase include: (a) higher cognitive functions (e.g. problem solving skills) (b) motor skills (e.g. colouring) (c) academic (e.g. sequencing) (d) communication (e.g. requesting) (d) interpersonal (e) learning readiness (e.g. imitation) (f) personal responsibility (e.g. advocacy) (g) play (e.g. playing with toys) (e.g. social and pretend play) (h) self-regulation (e.g. self-management). However, what is not clear in this review is whether the behaviour to increase were replacement behaviours based on the function of problem behaviour or a functional skill in isolation of a problem behaviour.. In addition, behaviours to decrease included: (a) problem behaviours (e.g. self-injury, aggression, disruption, destruction of property, or hazardous or sexually inappropriate behaviors) (b) sensory or emotional regulation (e.g. sleep disturbance) (c) restricted, repetitive, non-functional patterns of behavior, interests, or activity. Finally, the identification of a target behaviour informs the measurement method to be used and focuses the assessments that will be utilized to develop a behavioural intervention plan.

Multicomponent Treatment

The outcome of an FBA is a functional hypothesis (e.g. direct) or verified function (e.g. functional analysis) that informs the development of a behavioural intervention plan. Stated early in this review, determining the function(s) for the problem behaviours is integral to the development of an effective behavioural intervention (Sturmey, 1994). According to Smith

(2011) behavioural interventions to treat problem behaviour typically involves the manipulation or management of environmental events or conditions with the objective that the problem behaviour will be eliminated or reduced. Behavioural intervention components can be further categorized into a) antecedent manipulations b) skill development c) intervention (consequence) (Clarke, Dunlap & Stichter, 2002). Clarke et al. (2002) conducted a review of articles published on intervention research with children and youth with emotional behaviour disorder from 1980 to 1999. The results indicated that the majority of interventions were consequence based, followed by antecedent manipulations and skill development. In the review, multicomponent interventions averaged approximately one third of articles. Similarly, the National Autism Center (2015) found that the behavioural interventions utilized in the sample of 453 articles reviewed, included multicomponent interventions that utilized both an antecedent and consequence strategies. In this review, consequence strategies included both strategies to reduce problem behaviour and strategies to teach functional replacement behaviours. A study by Montgomery et al., (2013) revealed 7 common components in intervention plans in educational environments. These components included (1) Stimulus-Based Procedures (2) Instruction-Based Procedures (3) Extinction-Based Procedures (4) Reinforcement-Based Procedures (5) Punishment-Based Procedures (6) Systems Change (7) Unclear/Other. In addition, multicomponent interventions utilizing prevention, skill building and interventions sections have been demonstrated to be effective in reducing problem behaviour of children and youth with ASD in applied settings (Carr & Carlson 1993; Feldman et al., 2002, Horner et al., 2002). Furthermore, the following will provide an overview of each component of a behavioural intervention plan.

Prevention (antecedent manipulation). According to Smith (2011), antecedent manipulations involves the manipulation of antecedent events that (a) reliably occasions a

problem behaviour (e.g. function based) or (b) does not depend on identifying antecedent that reliable occasions a problem behaviour (e.g. default). Furthermore, Horner et al. (2002) recommended behavioural intervention plans include antecedent interventions that (a) emphasizes control of stimulus-based events that make the problem behaviour irrelevant and or altering an aversive stimuli. More specifically, examples of antecedent interventions include non-contingent reinforcement (Hagopian, Fisher, & Legacy, 1994), matched stimuli to establish and abolish motivation of problem behaviour (Rapp, 2007), altering an aversive stimuli (Cameron, Ainsleigh, & Bird, 1992), choice making (Carr & Carlson, 1993), embedding activities (Carr & Carlson, 1993).

Skill development. Horner et al., (2002) recommended developing behavioural intervention that teaches socially appropriate behaviours or replacement behaviour training based on the function of the problem behaviour. The underlying premise is that teaching an appropriate behaviour which serves the same function as the problem behaviour, and requires less effort, can replace the problem behaviour (Horner et al., 2002). Examples of replacement behaviours include functional communication training (Durand & Carr, 1991; Hanley, Iwata & Thompson, 2001), cooperation (Ducharme, 1996), self-management (Stahmer & Schriebhman 1992), and waiting skills that comprise building tolerance for delayed reinforcement (Carr & Carlson, 1993).

Intervention (consequence based). In order to decrease problem behaviour, behavioural intervention or behaviour management strategies typically include the organization of consequences to prevent reinforcement of problem behavior and/or the organization of consequences to maximize reinforcement of competing, appropriate behaviors (Horner et al., 2002). Interventions can also focus on decreasing problem behaviour through punishment (e.g., restitution). Examples of intervention strategies include: extinction (Fisher et al. 1998),

punishment (Hagopian et al. 1998), and reinforcer attenuation (Athens & Vollmar, 2010). Multicomponent treatment packages can be implemented through two approaches during service delivery (a) direct where intervention is provided by a therapist or (b) indirect where caregivers are trained to provide intervention.

Types of Service Delivery

In the province of Ontario, services based on ABA for children with ASD are provided through (1) Early Intensive Behaviour Intervention (EIBI) and (2) Mediator model or caregiver implemented intervention. EIBI can be characterized as a) intervention is individualized; b) behaviour analytic procedures are utilized to build new skills and reduce problem behaviours; c) normal developmental sequences guides selection of intervention goals and short-term objectives; d) intervention is typically delivered in one to one adult to children ratio then gradual transitions to small-group etc.; e) intervention typically begins in home and is carried into other environments; f) programming is intensive, includes 20 to 40 hrs; g) most children start intervention in preschool years 3 to 4 years of age (Eldevik et al., 2009; Reichow, 2011). Furthermore, positive outcomes have been associated with EIBI and children with ASD (e.g. Lovaas, 1987; Reichow, 2011; Virtues-Ortega, 2010).

In addition, Perry et al. (2008) conducted an "effectiveness" study (does it work in the real world?), with a sample of 332 children with ASD aged 2-7 years participating in a community-based publically funded EIBI program in the province of Ontario. Results indicated statistically significant and clinically meaningful improvements such as reduction in autism severity, gains in cognitive and adaptive levels, as well as doubling of children's rate of development and 11% achieved average functioning. Primary recipients of intervention in this model are the children, while the caregivers are involved in some capacity (i.e. generalization).

In addition to EIBI, the province of Ontario has also invested funding for services for children with ASD that involves the caregivers as the primary change agent or implementer of intervention.

The second type of service delivery model provided to families and their children with ASD, is referred to as a mediator model. Unless specified, the words “mediators” and “caregivers” will be used interchangeably to include a combination of parents, siblings, extended family, school employees (e.g., teacher, educational assistant), respite workers and others throughout this study. According to Feldman et al. (2002) a mediator model occurs when the professional visits the client on a weekly (more or less intensity based on needs) in their natural settings (e.g., home, day care, school) and provides training to the natural mediators to assist with FBA’s and implement intervention plans. More specifically, the mediators are trained by the professional to implement intervention plans to their child (if parents) or clients (if school personnel or group home staff). Research of mediator or parent implemented intervention has demonstrated that caregivers can successfully implement behavioural intervention plans and successfully reduce the problem behaviours exhibited by adults and children with diagnosed with ASD and intellectual disabilities (Feldman et al., 2002; Lucyshyn et al., 2007; Mclaughlin, Denney, Snyder, & Welsh, 2011; Meadan et al., 2009). Feldman et al. (2002) demonstrated the effectiveness of training caregivers in behavioural interventions to reduce problem behaviour and increase adaptive skills in individuals with intellectual and developmental disabilities (including ASD) in community settings. In addition, caregiver implemented interventions have also been demonstrated to be effective in the development of communication and social skills (Patterson et al., 2011), vocalizations, spontaneous request and play skills (Gillett & LeBlanc, 2007), feeding skills such as taste aversions (Sharp et al., 2014). More specifically, Patterson et al. (2011) found

in a systematic review of literature utilizing single subject designs large intervention effects across verbal language, child vocalization, child imitation and that parents of children with ASD were able to learn and implement strategies to increase their child's communication and social development over a short period of time. In addition, in a study conducted by Suess et al., (2014) positive outcomes were observed with the use of telehealth to teach parents of children with ASD functional communication training and a reduction in problem behaviour. Finally, Horner et al. (2002) found that behavioural interventions implemented by typical agents (parents, teachers) produced a greater magnitude of behaviour reduction compared to atypical agents (specialists) or settings.

Mediator Training

There are a variety of approaches to training mediators on how to implement behavioural intervention plans that include: (a) *in vivo* individual training; (b) group training; (c) support groups with an educational component; (d) self-directed utilizing training manuals and instructional videos (National Autism Center, 2015). Although several approaches to training caregivers are listed above, some of which involves enhancing knowledge or verbal skills, or using a didactic approach, training caregivers requires the focus of enhancing performance skills that is data based or that observations can be documented of caregiver skill acquisition (Parsons, Rollyson, & Reid, 2012). An approach to training mediators that is based on enhancing performance skills is referred to as Behaviour Skills Training (BST), that includes the following components (a) instructions (b) modeling (c) practice, and (d) feedback (Parsons et al., 2012; Miles & Wilder, 2009). BST has been demonstrated to be an evidence based training model for teach caregivers to implement discrete trial teaching (Sarokoff & Sturmey, 2004); utilize guided

compliance for non-compliant behaviour (Miles & Wilder, 2009); and teach requesting items/activities using incidental teaching (Hsieh, Wilder & Abellon, 2011).

Treatment and Implementation Integrity. Once a mediator has been trained to implement intervention, it is highly recommended that implementation be monitored to ensure that treatment integrity is maintained (Playnic, Fefferi, & Maupin, 2010). Treatment integrity refers to the degree to which an intervention plan is implemented as planned (Wilder, Atwell, & Wine, 2006). Moreover, research has demonstrated that high treatment integrity is associated with positive outcomes (Fiske, 2008; Wilder et al., 2006); reducing problem behaviour (DiGennaro, Martens & Kleinmann, 2007); increasing alternative skills (Plavnick et al. 2010). However, in applied settings, the implementation of behavioral intervention plans are less likely to have perfect integrity (Wilder et al. 2006). Nonetheless, it is recommended to utilize effective approaches such as direct training to caregivers to increase treatment integrity (Fiske, 2008).

As previously described, one responsibility of a behaviour consultant is train mediators to implement intervention plans (Feldman et al., 2002). Therefore, integrity checks can be extended to the type of strategies utilized by the consultant to train mediator's how to behaviour intervention plans. According to Fettig, Schultz & Sreckovic (2015) this type of integrity check is referred to as fidelity of implementation. More specifically, fidelity of implementation is the integrity check that provides an assessment of the degree to which the training conducted by the consultant (e.g. practitioners) with the mediators was as designed (Fettig et al., 2015). Finally, a high fidelity of implementation of behaviour intervention plans with the caregivers also increases the likelihood of positive outcomes (Fettig et al., 2015). Similarly to checks for treatment integrity and fidelity of implementation, program evaluations provide information on whether

activities performed in a service are consistent with what was promised and or expectations of that service.

Program Evaluations

Program evaluation is the process by which organizations determine the degree to which they meet their intended goals. More specifically, evaluators gather information and provide feedback on program outcomes to interested stakeholders (e.g. previous or future clients, funding sources, boards of governance); enable service providers to improve effectiveness of services; assist government officials to make program level decisions (Posavic, 2015). In addition, the results or output of an evaluation supports government agencies and private organizations to identify needs; facilitate effective planning and monitoring; accurate assessment of quality service; nurture improved practices and detect unwanted program side effects (Posavic, 2015). Therefore, the model that will be outlined below and utilized in this study was described by Posavic (2015). According to Posavic (2015), the goals of program evaluations can be met with four common types of program evaluations. The first type of program evaluation, "Assess needs of program participants", involves the measure and assessment of the program participants, community and or organization unmet needs. The second type of program evaluation, "Examining the process of meeting the needs", involves examining the process and service activities of meeting the needs of the program participants and or organization (e.g. type of services provided, the degree of how the program operated as expected; the nature of people served). The third type of program evaluation, "Measure the outcomes and impacts of programs", involves the measurement of outcomes of the program participants (e.g. Are the participants performing the expected skills during and after the program?). The final type, "Integrating needs, cost and outcomes", involves the measure of the program efficiencies and

cost effective analysis determined by combining the outcomes, costs and needs of the program. Although, these types of evaluations are outlined as separate entities, they can also be combined to provide an evaluation (Posavic, 2015). As will be highlighted below, types 1 through to 3 will be utilized in the development of the research questions for this thesis.

Program Evaluations for Children with ASD

The following section will provide an overview of published evaluations of community based services for children with ASD that followed a mediator model approach or similar (e.g. caregiver implement intervention). Heitzman-Powell et al. (2014) conducted an evaluation of an ABA outreach training program for parents of children with autism who resided in remote locations. Training was provided to 7 parents from 4 families via web-based and telemedicine technology that required support for their child. The program was evaluated by providing families with pre and post-test skills and knowledge assessments. The results revealed that knowledge gains were promising and all parents demonstrated substantial gains in implementation of ABA skills. Although an efficient way to provide services in remote geographical locations, there was no child outcomes reported. Baker-Ericzén, Stahmer & Burns (2007) evaluated a pivotal response training program of 158 families and their children with ASD. Parents were provided with training by therapists over a 12 week period. Outcomes were measured by utilizing the Vineland Adaptive Behaviour Scales pre and post treatment. The results indicated that the participants improved significantly in adaptive functioning at the end of the 12 week period. These studies provide an illustration of the methods commonly used to evaluate programs for children with ASD, namely indirect and or standardized measures. Although, there are benefits to using these measures (e.g. efficient, psychometric properties; Burns 2014) what is not captured is the clinical data that are utilized in these applied settings.

More specifically, results from a standardized measure may indicate statistically significance of utilizing an intervention, however, it becomes difficult to determine whether clinically significant improvements have occurred for individual participants due to the intervention (Roane et al., 2011). Data that are utilized on a daily basis by the practitioner to determine the course of treatment and or treatment effectiveness is not captured in indirect/standardized measures commonly used in program evaluations.

Similarly, other program evaluations report outcomes using indirect measures and in the absence of child outcomes. Dyer et al. (2006) reported high quality of life scores by caregivers during a program evaluation of services for children with ASD. Sharp et al. (2014) evaluated a parent-training curriculum for children with ASD who experienced taste aversions and low intake diet. The results determined high social validity from parents for the treatment, but, there were no child outcomes.

The type of outcome data across these studies varied from standardized assessments of child outcome (Baker-Ericzén, Stahmer & Burns, 2007) to measurement of caregiver knowledge (Heitzman-Powell et al., 2014) but did not include clinical data (i.e., direct measurement by consultants or caregivers of the specific behaviours targeted for change). Furthermore, a review of the literature revealed that the use of clinical data in treatment evaluations is very rare with children with ASD. However, clinical data or single subject data were utilized by Burns (2014), in a program evaluation of children with emotional and behavioural disorders at a mental health centre. The single subject data of the 95 participants (approximate 10% of sample children with ASD) were aggregated and grouped into two constructs (a) improved behaviour from month to month (for “improved” an indicator of progress or replacement behaviour must have increase by 5% from the previous month and a referral behaviour such as aggression must have decreased at

any amount), (b) maintenance of behaviour at optimal levels (indicators of progress is 95% from one month to the next and no incidences of referral behaviour from 1 month to the next). With these constructs, Burns (2014) found that youth demonstrated improvement from one year to the next for indicators of improvement and maintained these behaviours at optimal levels. According to Burns (2014), traditional pre/post methods used in program evaluations may not capture the individual features of interventions that contribute to the outcomes compared to single subject designs. The aforementioned review highlights the lack of research of community based or "effectiveness" evaluation of mediator or parent implemented intervention for children with ASD.

Rationale

As the prevalence of children and youth diagnosed with ASD increases, government officials continue to invest in services and supports for this population. ABA-based services are funded by MCYS to provide services based on applied behaviour analysis to children and youth with ASD and their families. Provided that services are based on ABA, evaluators of these programs need to examine the congruency of the actual service activities to best practices highlighted by both the BACB and in the literature on measurement methods, assessment and intervention strategies. In addition, as highlighted in the literature review, program evaluations of community-based services for children with ASD have excluded the use of single subject data in outcome measures. Similarly, a dearth of literature exists of program evaluations of services for children with ASD who exhibit problem behaviour with the exception of Suess et al (2014) and Feldman et al., (2002), who included some children with ASD in their sample of participants with developmental disabilities. In addition, the inclusion of single-subject data utilized during intervention for problem behaviour would allow for computations of intervention effects.

Therefore, an evaluation of outcomes utilizing clinical data would be both beneficial to stakeholders of the ABAS being examined and would contribute to the literature on the effectiveness of a caregiver implemented behaviour intervention to reduce problem behaviour exhibited by children with ASD. Information about the assessment and treatment methodologies used and their outcomes can contribute to the evidence of effectiveness of ABA in real settings. It can contribute knowledge to ABAS stakeholders about the strengths and limitations of the clinical processes and outcomes.

Specifically, this study was designed to answer a series of questions informed by the different components of program evaluation outlined by Posavic (2015) including (a) assessment of participant needs, (b) examination of process and service activities, and (c) measurement of outcomes and impact. Examination of the cost effectiveness could not be undertaken due to the retrospective design of the study.

Research Questions

The following research questions will be the focus of this study.

1. What are the presenting problems of the sample targeted during intervention?

The first type of evaluation outlined by Posavic (2015) was to assess the needs of the program participants. As described above, all the participants in this sample were referred to ABAS for concerns with problem behaviour. The presenting problems or needs of the sample that will be examined will include (a) the type of problem behaviour(s) targeted during intervention of the total sample and across behaviour plans with measurement focused on replacement behaviour and behaviour plans with measurement focused on problem behaviour. A further description of these two groups will be provided in the following sections.

2. To what extent were service activities and processes consistent with the expectations outlined by professional, organizational, and governmental standards in this sample?

The second evaluation type outlined by Posavic (2015) was to examine the service activities and process that were utilized to obtain the outcomes achieved. Therefore, the service activities and process undertaken by the ABAS will be examined to determine the agreement implementation fidelity (Fettig et al., 2015) with (1) expectations outlined by the MCYS for ABA-based services (2) service and professional standards. More specifically, service activities such as service recipients, duration, context and mediators trained will be compared to the expectations outlined by MCYS, whereas, assessments, treatment and service delivery activities will be compared to the clinical standards utilized by the organization. Therefore, the first hypothesis is that the services activities and process utilized by the consultants will follow the standards outlined by all three groups.

3. Are there meaningful treatment outcomes at the conclusion of services?

The third evaluation type outlined by Posavic (2015) is measuring outcomes. In this evaluation, service outcomes will be measured by percentage of behaviour change and effect size computations for both groups (a) plans with measurement focused on problem behaviour and (b) plans with measurement focused on replacement behaviour. The third hypothesis is that meaningful improvements will be demonstrated across both percentage change data and effect sizes.

Method

The method section begins with a description of the ABAS program as it existed at the time of the service delivery evaluated in this study. Following the service description, is a description of the specific methods used to carry out this retrospective program evaluation study.

Program Description

The Applied Behaviour Analysis Service (ABAS) is a collaborative initiative provided by four community agencies, with funding to provide services to families of approximately 1000 children and youth with ASD each year. During the time referenced in this study, referrals were made by the primary caregiver and, in some cases, by the child/teen in circumstances whereby they had the ability to do so. Referral areas included problem behaviours or specific areas of skill deficits. Eligibility required proof of diagnosis of ASD. At the time of intake, an electronic record was developed for each new referral. Clinical information including the referral information, and reports from other clinicians, (e.g., speech pathologists, psychologists) was stored in this record and key documents related to service were included. Once referred, the client was placed on the waitlist for one of two streams: (a) clients requesting assistance primarily to develop skills in a particular domain (e.g., social skills, communication, activities of daily living) without reference to problem behaviour or (b) clients requesting assistance primarily to treat a problem behaviour. Eligible cases for services were activated on a first come first served basis based on the date of the referral and irrespective of problem behaviour.

The vast majority of services were delivered in the client's own home, with a minority of services delivered in other community settings, which included schools, community centres and/or other relevant locations. Services were provided by behaviour consultants whose roles consisted of (a) introducing the service to families (b) conducting and developing assessments (c) developing the intervention plans in conjunction with the family (d) Training the caregivers to implement the intervention plans. The majority of services were based on a mediator consultation model, in which the child's caregivers were trained to implement the intervention. Less frequently, treatment was delivered directly through a group format with either children

and/or their parents, however these data are not included in this study. Service consisted of an initial FBA for problem behaviour referrals and a functional skills assessment for adaptive behaviour referrals, followed by weekly average 1 to 2 hours per week of consultation for approximately 2-3 months. A one month follow up meeting was conducted with most of the caregivers in the treatment setting and in some instances, at the request of the caregiver, follow up was completed by phone.

Goal setting. At the onset of service delivery for participants whose primary target is the treatment of problem behaviour, a functional behavioural assessment is conducted by the consultant. Once the assessment is complete, a single goal is developed with the caregivers in collaboration with the consultant using a SMART approach: Specific, Measureable, Attainable, Realistic and Time bound (Lawlor & Hornyak, 2012) and a measurement system is designed to collect clinical data to determine the effectiveness of the intervention plan designed to meet the goal. The goals developed for the behaviour intervention plans for those referred for treatment of problem behaviour is inherently to decrease problem behaviour. However, the measurement system designed by the consultant can focus on (a) collecting data to monitor the problem behaviour intended to decrease or (b) collecting clinical data to monitoring the functional replacement skill intended to increase. If the goal specified by the caregiver focuses on the decrease in problem behaviour, the graph summarizes clinical data related to the problem behaviour. If the goal specified by the caregiver focuses on increasing a skill intended to replace the problem behaviour, the graph summarizes the clinical data related to the replacement skill being taught, and may or may not include a separate trend line for the problem behaviour. Though both types of intervention have an overarching goal to decrease problem behaviour,

formal monitoring of the problem behaviour is not always included when the focus of the SMART goal to increase a functional replacement skill.

Functional behaviour assessment. The FBA process includes multiple descriptive methods of assessment. The consultant in collaboration with the caregivers and the clients, when applicable, complete the FBA. The methods include the use of semi-structured interviews developed by the program. The behavioural contextual interview includes questions about possible antecedent and consequent influences on the target behaviour i.e. what happens before the problem behaviour (Cooper, Heron & Heward, 2007). The Bio-Psycho-Social interview includes questions aimed to gather medical (i.e. health, medications) psychological (i.e. skill deficits) and social (i.e. school status, difficulties in social settings, relationships) (Griffiths & Gardner, 2002). Other indirect methods available include questionnaires such as Questions About Behaviour Function (QABF; Matson & Vollmer, 1995) and Functional Assessment Screening Tool (FAST; Iwata & DeLeon, 2005). Direct methods available include antecedent, behaviour, consequence, (ABC) recording (Bijou, Peterson & Ault, 1968) by the consultant and/or caregiver. Finally, experimental methods such as structural and functional analyses can be utilized by the consultant (Hanley, Iwata & McCord, 2003). Given the short duration of service, the FBA can take between 2-3 hours of direct hours and is expected to be completed within 2-3 weeks of case activation. Once the functional behavioural assessment is complete, the consultant is expected to formulate a functional hypothesis based on the indirect and descriptive methods or a verified function based experimental methods. This functional hypothesis is integral to development of the treatment plan.

Program development and implementation. As described above, behavioural treatment plans were developed by the consultant with the support of a supervising Board Certified

Behaviour Analyst (BCBA), and input from the caregivers. A BCBA is a professional who has obtained a graduate-level certification in behavior analysis, required to comply with the Behaviour Analyst Certification Board (BACB) ethics requirements and supervises others who implement behavior-analytic interventions (BACB, 2016). The BCBAs in the ABAS review, revise as needed, and sign off on all behavioural treatment plans. The principal student investigator of this study is a supervising BCBA in this organization. Treatment plans are expected to include strategies to prevent problem behaviour (e.g. non-contingent reinforcement (Hagopian, Fisher & Legacy, 1994), to develop replacement behaviours (e.g. functional communication (Durand & Carr, 1991), and to intervene when problem behaviour occurs. Behaviour Consultants, with support from Behaviour Technicians, train the mediators to implement the treatment plan by using a behavioural skills training (Parsons, Rollyson, & Reid, 2012) or a general case training (Ducharme & Feldman, 1992) approach. The training sessions include the consultant modelling the strategy to be taught, practicing with the caregiver, observing the caregiver implementing the strategy, and providing performance feedback related specifically to the caregiver's implementation of the treatment plan.

Service delivery standards. Utilizing the evidence in the literature and the standards outlined by the BACB, service delivery standards were developed by the ABAS BCBAs in collaboration with the consulting psychologists. The service delivery expectations are broken down into three categories: (1) Assessment (2) Treatment Components (3) Mediator training. Not included in the list below, however a standard in the ABAS, is that all service activities must follow the professional and ethical compliance code developed by the BACB (e.g. consent, least to most intrusive procedures). In addition, refer to Table 6 below for the reliability for the required items as per the service delivery standards. Note that reliability for the categories for the

specific treatment components is included, but the reliability for specific treatment procedures was not adequate for use in this study (Awan, 2016).

(1) Assessment:

a) A problem behaviour must be identified and measured

b) A FBA must be conducted prior to the development of an intervention plan. At a minimal, a functional behavioural assessment needs to include at least two methods (e.g. two of indirect/direct/experimental)

c) A functional hypothesis must be identified (or identified function if a functional analysis is conducted)

(2) Treatment Components: Intervention Behaviour plans must include the following sections and strategies (i) Prevention (ii) Skill Building or replacement behaviour (iii) Behaviour Management

(3) Service Delivery: A Behaviour Skills Training approach or components of BST are utilized when training the mediator(s) to implement behavioural intervention plans.

Study Description

Inclusion criteria. As this was a retrospective file review study, participants were not recruited directly. Children and adolescents who were served by ABAS will be referred to as "participants" in this study. Further, Research Ethics Board Clearance was obtained at from the Social Sciences Research Ethics Board at Brock University and from the Research Ethics Board at ABAS. The criteria for inclusion in the study were: (a) service delivery window from November 1 2013 to July 30 2015 (b) evidence of service provided (e.g. goal met or goal not met); (c) target behaviour and goals pertaining to the behaviour domain (e.g. referrals for

reducing problem behaviour); (d) minimum of three daily data points at baseline and three data points at treatment were available and graphed; (e) key documents were available in the file including the behavioural intervention plan, the FBA and service provision plan. The service delivery model required that these documents be completed, so any files excluded due to lack of documentation were likely due to a lack of consistent storage and maintenance requirements. A thorough description of the screening process is provided in the procedure section below. Based on these criteria, 112 children and youth with ASD who had previously received service from ABAS were identified for inclusion in the study.

Measures

1) The Behaviour Assessment Treatment Taxonomy (BATT; Condillac, 2009d). The Behavioural Assessment and Treatment Taxonomy (BATT) was designed as a systematic monitoring or file review tool to track assessment, treatment, and measurement methodologies used, and to summarize results of assessment, and intervention from single-case designs used in clinical practice (see Appendix A). The BATT was intended for prospective use by the behaviour consultant during the course of service provision. In this study, the BATT was utilized retrospectively to systematically retrieve information from the clinical files.

Content. The BATT is divided into five sections including 1) client demographics and target behaviour 2) behavioural assessment 3) behavioural intervention strategies 4) data collection, and 5) case summary section. Please refer to Table 1 below for a description of each section. Finally, the BATT allows for the addition of agency specific variables of interest in the "other" category in each section such as specific assessment methods or measures not indicated on the BATT.

Table 1

BATT Description of Sections

Section	Description
1) Client Demographics and Target Behaviour	The first section, captures non-identifying client demographic information such as age and the problem and replacement behaviour(s) targeted in the intervention plans.
2) Behavioural Assessment	The assessment section captures current school status, medical status, skills assessments such as a functional communication record, indirect preference assessments, and functional behavioural assessments methods. Further, the functional behavioural assessment subcategory includes both (a) indirect (e.g. interviews, questionnaires) and (b) direct descriptive assessments (e.g. ABC incident recording, informal probes); (c) experimental methods (e.g. structural, functional analysis); (d) function(s) utilized in programming
3) Behavioural Intervention Strategies	The treatment section captures the types of treatment components utilized such as antecedent strategies, teaching adaptive behaviours/skill building, and consequence strategies, and the type of training utilized with the mediators (e.g. feedback, role-playing).*
4) Data Collection	The data collection section documents the measurement methods used by the caregiver after the behavioural plan is introduced (e.g. frequency, duration)
5) Case Summary Section	The case summary section includes the (a) goal of the treatment plan (b) direction of desired treatment effect(s) including increase(s) in skills and/or decrease(s) in problem behaviour (c) the specific type of data represented on the graph (e.g., frequency, percentage, duration, latency, intensity); (d) 3-5 data points in baseline, treatment and follow-up data points.

* Please note: Issues with reliability (presented below) precluded the use of data of specific treatment components.

Reliability. The BATT was utilized to extract and code information from the files reviewed for this study. A description of both the extraction and the abstraction phases will be reviewed in the procedure section below. In addition, a companion thesis (Awan, 2016)

examined the inter-rater reliability of the BATT in this sample. The student principal investigator on the reliability study completed data extraction for 50 files for reliability purposes (Awan, 2016). The files for the reliability study were randomly selected by the student investigator conducting the reliability study and were evenly distributed across the three extractors whereby 46% of files extracted by each person were examined for reliability purposes. Mean reliability across all raters for the extraction phase was excellent (97% agreement across 50 files) as calculated by percentage agreement. As found in Awan (2016), percent agreement was calculated for a total of 71 items extracted per file. Specific details on the reliability for the extracted items are included in Table 4. More specifically, high percent agreement was observed for target behaviours, assessment methods, treatment components and data points that are utilized in this evaluation. Similarly, there was excellent mean reliability for the abstraction phase (*kappa* >.80, p<.001) for the coding of replacement and problematic target behaviours (*kappa* >.80, p<.001; see Table 2 and 3). As seen in tables 5 and 6, there are fewer replacement behaviour categories (3) compared to behaviour categories (7). All results reported in this study are reported at the categorical and/or item level with acceptable reliability scores.

Table 2

Percent Agreement of Items for BATT across Coders

Items	Percentage
Target Behaviour to increase	84%
Target Behaviour to decrease	92%
Current School Status	100%
Current Medical Status	100%
Functional Communication Record	98%
Indirect Preference Assessment	100%
Behaviour Contextual Interview	98%
Bio Psycho Social Interview	98%
FBA Question QABF	94%
FBA Question FAST	100%
FBA Direct ABC Consultant	94%
FBA Direct ABC Caregiver	74%

FBA Direct ABC Unspecified	88%
FBA Direct Informal In vivo Probe	80%
FBA Direct Other	96%
Overall Hypothesized Function	90%
Prevention Strategies	94%
Skill Development Strategies	94%
Intervention Strategies	98%
Case Summary Direction Effect	90%
Case Summary Type Data Collected	100%
Case Summary Baseline Data Point	92%
Case Summary Treatment Data Point	90%
Number of Recommendations	88%

Source: Awan, 2016.

Table 3

Reliability Data for Problem Behaviour

Item Abstracted	N	Percent Agreement	κ (95% CI)*	PI	BI
Dangerous behaviours	22	99.17%	0.97 (0.92-1.00)	0.63	0.01
Aggression	18				
Self-Injury	3				
Destruction	1				
Inappropriate Communication	20	97.52%	0.91 (0.81-1.00)	0.68	0.01
Aggressive Threats	2				
Swearing	3				
Name Calling	9				
Repeated Questions	6				
Refusal/Protest	14	99.17%	0.96 (0.89-1.00)	0.76	0.01
Disruptive – Crying	12	100%	1.00 (1.00-1.00)	0.80	0.00
Disruptive – Screaming	16	99.17%	0.96 (0.89-1.00)	0.74	0.01
Disruptive – Other	13	100%	1.00 (1.00-1.00)	0.79	0.00
Whining	5				
Tantrums	4				
Throwing	4				
Miscellaneous Behaviours	24	98.35%	0.95 (0.87-1.00)	0.62	0.02
Taking or Grabbing Items	3				
Flopping	2				
Running	7				
Verbal Stereotypy	1				

Other	11

Note. κ = Cohen's Kappa; CI = Confidence Interval; PI = Prevalence Index; BI = Bias Index.
* $p < .001$
(Awan, 2016)

Table 4

Reliability Data for Replacement Behaviours

Item Abstracted	N	Percent Agreement	κ (95% CI)*	PI	BI
Cooperation	11	100%	1.00 (1.00-1.00)	0.53	0.00
Tolerance and Waiting	10	97.87%	0.93 (0.81-1.00)	0.60	0.02
Tolerance	5				
Waiting	5				
Other Adaptive Behaviours	26	93.62%	0.87 (0.73-1.00)	0.09	0.02
Transitions	2				
Routine/Schedule	3				
Sharing	1				
Play Skills	4				
Activities of Daily Living (ADL)	3				
Safety Skills	2				
Self-Management	1				
Communication Mands	3				
Communication Tacts	1				
Communication-Intraverbals	2				
Other	4				

Note. κ = Cohen's Kappa; CI = Confidence Interval; PI = Prevalence Index; BI = Bias Index.
* $p < .001$
(Awan, 2016)

Similarly, as highlighted in Table 7, high percent agreement (93%) was obtained across the four extractors of overall agency required items. The items bolded in the chart represent the required items that are expectations of the organizational and that will be further explained below.

Table 5

Sample Size and Percent Agreement for Required Extracted Items with Occurrence Data

Extracted Item	N	Percent Agreement
(1) Target Behaviours – Overall	**100**	**88%**
(2) Functional Behaviour Assessments – Overall	**900**	**96%**
(3) Overall Hypothesized Function	**69**	**90%**
(4) Treatment Recommendations – Overall	**150**	**95%**
Prevention Strategies	50	94%
Skill Building Strategies	50	94%
Intervention or behaviour management Strategies	50	98%
(5) Mediator Training – Overall	**404**	**95%**
BST Model	50	96%
Instruction	50	90%
Modeling – Live	50	96%
Modeling – Video	50	98%
Roleplay	50	92%
Feedback	50	92%
Other Mediator Training	50	100%
(14) Required Agency Items – Overall	**1496**	**93%**

(Awan, 2016)

2) The Case Study Behaviour Categories by Atkinson & Feldman (1994)

Problem Behaviours. The Case Study Behaviour Categories by Atkinson & Feldman (1994) was adopted to develop the coding categories for the problem behaviours. In addition, the level of percent agreement and reliability of coding target behaviours (see Table 3 above) also informed the development of these categories.

Table 6

Case Study Behaviour Categories: Description Problem Behaviours

Section	**Description**
(1) Dangerous Behaviours	Includes any one of the following: a) Aggression to others: (but is prevented or misses) or actually hits, slaps, punches, bites, pinches, scratches, pokes, kicks, shoves, or throws objects at another person with sufficient intensity to inflict or potentially inflict immediate pain and/or injury to the victim b) Self-injury: attempts to (but is blocked) or actually hits, slaps, punches, bites, pinches, scratches, pokes, kicks own body, or non-accidentally brings body part in contact with hard object with sufficient intensity to cause immediate accumulated injury c) Property Destruction: attempts to (but is prevented) or actually causes property damage.
(2) Dangerous Behaviours plus other Problem Behaviour	Any combination of the Dangerous behaviours listed above with another problem behaviour (e.g. refusal). Please note that "other problem behaviour" in this category does not include disruptive behaviour.
(3) Inappropriate Communication	Includes any one of the following: a) Aggressive Threats- verbally or nonverbally (e.g., raises fist) threatens to harm another person b) Self-talk – speaks words no directed at anyone else c) Aberrant Speech – bizarre content and/or inflections d) Swearing: the use of profanities e) Name calling: the use of inappropriate words directed to others f) Repeated questions: asking a question(s) appropriate to context and directed to someone but not at appropriate rate.
(4) Refusal:	Includes any one of the following: a) Verbal or physical behaviour directed at someone (shaking head, pushing items away, ignore), any mention of being non-compliant/not cooperating, ignore request b) Non-Specified Behaviour: i.e. Non-compliant c) Off task behaviour: child is not engaged in a task (i.e. academic work)
(5) Disruptive Behaviour	Includes any one of the following: a) Screaming plus crying/whining: Includes any one of the following: (i) Screaming: Vocalizes loudly or loud tone, could include yelling (ii) Crying: Weeping (with tears) (i) Whining: making a loud pitch, complaining sound (cry or vocalization). b) Tantrum/throwing: Includes any one of the following: (i) Tantrums: stomping feet, slamming doors (ii) Throwing: throwing items

(6) Miscellaneous: Includes any one of the following:

- (i) Physical Stereotypy: rocking, finger flicking, head-waving, spinning objects, twirling self, constant touching
- (ii) Verbal Stereotypy: includes vocalizations that are not related to the context (non-contextual vocalizations)
- (iii) Stripping: Removing clothes in public where not socially appropriate
- (iv) Inappropriate touch others: Attempts to or succeeds at making physical contact with another person with was not invited, expected or socially appropriate
- (v) Public Masturbation: touches, rubs, and/or exposes genital area in a public setting
- (vi) Taking or grabbing items: Taking items/food from others in proximity to them
- (vii) Flopping: Any part of body falling to the ground
- (viii) Running: Moving faster than a walking pace in a situation where doing so is not expected and could be disruptive or dangerous
- (ix) Fecal Smearing: Feces is smeared on self, others, environment
- (x) Urine/Bowel Accident: Urinating and/or defecating anywhere but in the toilet of diaper
- (xi) Pica: attempts to (but is blocked) or actually mouths a non-nutritive substance or object
- (xii) Vomiting/Regurgitation: Brings up previously swallowed food
- (xiii) Other: Any other socially appropriate inappropriate behaviours observed, but not listed above such as covering eyes, clench fists/swing legs, yawning, spilling food from plate, and getting out of bed.

Replacement behaviours. The Case Study Behaviour Categories by Atkinson & Feldman (1994), was adapted to develop the coding categories for the replacement behaviours (See Table 7). In addition, the level of percent agreement and reliability for coding the replacement behaviours (see table 4) indicated reliability only across three behaviours and or combinations. As a result, all replacement behaviours utilized in this study were included in one of the aforementioned behaviour(s).

Table 7

Case Study Behaviour Categories: Description Replacement Behaviours

Replacement Behaviour	**Description**
1) Cooperation	If written as "cooperation" "completing" "on task' or to exhibit a response to an instruction.
2) Tolerance/Waiting	Includes any of the following: a) Tolerance of non-harmful non-preferred routine activities/people/noises, non-harmful routine items: Word is written in definition, accepting responses from others, changes to routine/environment/schedule, "accept" b) Waiting for preferred routine/activities/people/noises of explicitly written as "waiting".
3) Other Adaptive Behaviour	Includes any of the following: a) Transition/Routine: (i) Transitions: Word is written in definition (ii) (ii) Routine: includes teaching how to follow a routine, visual schedule b) (B) Sharing/Play skills: (i) Sharing: sharing toys, turn taking (ii) Play skills includes teaching a specific play skills, parallel play, board game. c) Activities of Daily Living (ADL) /Homework/Safety Skills (i) ADL: includes trying new foods, toileting, showering, dressing/undressing, purchasing items or engaging in activities (ii) Homework: includes teaching how to complete homework, specific academic tasks i.e. worksheets (iii) Safety skills: crossing the street, looking both ways, holding hands during walk, staying within close proximity to caregiver during outings in the community. d) Self-Regulation/Management and other: (i) Self-Management: includes self-monitoring, self-instruction, habit reversal (ii) Self-Regulation: relaxation deep breath, progressive relaxation, counting (iii) Other adaptive behaviour: Not specified. e) Communication: (i) Mands (requests), includes spontaneous mands, request (ii) Tact (labeling): Teaching the child to label i.e. feeling (iii) Interverbals (Asking/responding to question, conversations) (iv) Communication other-not specified could include: Functional communication, f) Communication and other function skills: Includes any target within the communication category above and any of the other coding categories (e.g. ADL).

Procedure

Training and qualifications of extractors/coders. Four extractors and coders were involved in the process described below. The student principal investigator completed graduate training in master's program specializing in applied behaviour analysis, and is a BCBA. The student principal investigator has a dual role in this research as he is a supervisor in the ABAS. Three research assistants in total were involved in this process. Please note, two of three research assistant are current behaviour consultants in the ABAS, and were not provided with their own cases to extract/code. Finally, the third research assistant was a student in Dr. Condillac's lab. In addition, the principal student investigator in the companion project and the other 2 assistants have received and or completing graduate training in a master's degree specializing in applied behaviour analysis.

Screening. With appropriate permissions through the University and Agency Research Ethics Boards, a multi-step screen process was conducted to maximize inclusion for review while minimizing risks (e.g. access to personally identifying information). At the onset of the screening process, a sample of 200 cases was the goal to be included in this evaluation. An initial service window from April 1, 2014 to July 30, 2015 was determined as the starting point to screen cases using steps 1 and 2 listed below.

1. The first screening involved the student principal investigator reviewing all consultant caseloads during the designated time frame through the agency's secure electronic records. A master list of client case numbers, unique study identifiers and non-identifying demographic information were assembled and maintained on the secure network by the student principal investigator of this study. The student principal investigator reviewed all files to determine if the

final cases status was recorded as either goal met or unmet. Those excluded in the initial screen had a final case status of no service or no contact or fell outside of the service ranges.

2. A second screening by the principal student investigator involved reviewing individual client files on the secured electronic record that resulted in the identification of cases who received service between April 1, 2014 to July 30, 2015. During the second screening, cases were included if they met all of the following screening criteria:

a) if a problem behaviour was identified as the target of treatment: Indicated by treatment plan stating "goal is to decrease a problem behaviour" and or "goal is to increase replacement behaviour in absence of problem behaviour"

b) if services occurred in the target date range indicated in inclusion criteria

The principal student investigator repeated step 1 and step 2 of screening by going back one month at a time (e.g. from April 1 2014 to March 1 2014 etc.), until a sample of approximately 200 cases were identified to be eligible for this evaluation. As a result, 219 out of a sample of 482 cases were identified to be eligible based on the aforementioned screening steps.

Please note that for the first and second steps of screening it was not possible to collect reliability data as access to the caseloads and electronic records was limited given the sensitive nature of the information included in the full record.

3. The third screening was conducted by both the principal student investigator and the student from the companion thesis and involved locating the physical files and the key documents for the 219 cases from secured cabinets at the agency locations, which they both had consent to access. This resulted in a total of 207 cases that were identified based on the availability of key required documents including (a) the behavioural intervention plan, (b) the service provision plan, (c) the

FBA, (d) a minimum 3 baseline points and 3 treatment points either on a graph or on loose data sheets.

4. The final screening conducted by the principal student investigator and the student from the companion thesis identified the final sample of 112 based on the type of data that was available in the file. In this final screening, 95 of the 207 files were deemed ineligible because the available data was graphed as a weekly average, but there was not adequate information in the file to determine if a week included seven days or five days, to equalize for inclusion with the daily data. The daily clinical data from the eligible sample was measured across any one of the following dimensions frequency, duration, latency, intensity and percentage.

In summary, A sample of 188 participants who received services for behaviour concerns were excluded from this study during the screening process. The age range of the excluded clients was 3 to 18 years old (M=9, SD=3.48). The sample included mostly (87%) boys and fewer (13%) girl. The length of services ranges from 3 to 22 weeks (M =11, SD=3.5). Similar to the eligible sample, the majority of the sample received services for the first time (61%). The reasons for the exclusion of these participants include: (1) 47% Limited or no data available in the file (2) 53% only weekly data points available. Please note that there was a time in service that raw data was deleted once summarized in the service provision report.

Extraction. Once the screening process was completed, the BATT was utilized to extract file information from the final sample of 112 files. As described in Awan (2016), across three consecutive test files, the extractors achieved 90% agreement on the extracted information summarized on the BATT. Data extraction was completed by the student principal investigator for 73% of files, and two research assistants each extracted 13% of files. The student from the companion thesis extracted reliability data from 50 randomly selected files.

Abstraction. Once extraction was completed, reliability on the coding scheme for the target behaviours and the treatment components was completed to a criterion of four consecutive files coded to over 90% agreement. Once agreement was established in training, the target behaviours (e.g. problem and replacement), and treatment components (e.g. prevention, skill development, intervention) of behavioural intervention plans for n=112, were coded. Similar to the extraction phase, the student principal investigator coded 73% of files, and the two research assistants each coded 13% of files. The student from the companion thesis coded reliability from 50 randomly selected files. Once the abstraction phase was complete, all the data from the raters were pooled and entered into SPSS statistical software version 22.0 for analysis.

Outcome Data. The outcome data used in this study was extracted from the graphs or raw data found in the participants' clinical files. While, all intervention plans were designed to reduce problem behaviour, the data collected for each participant were based on either measurement of the acquisition or use of replacement behaviour (in the absence of the problem behaviour) or on the measurement of the occurrence of the problem behaviour. Thus two different successful data trends were possible. For ease of analysis the sample was subdivided into two groups: (1) those with data collection or measurement focused on the problem behaviour or (2) those with data collection or measurement focused on replacement behaviours. As a result, the presenting problems of the sample that were examined include: the type of problem behaviour(s) targeted for the entire sample and those that were the target of data collection (measurement).to achieve the SMART Goal. As this was a retrospective study, reliability on these data were limited to checking accuracy of extraction. Inter-observer reliability of the outcome data was not collected during service provision and is therefore unavailable for this study.

Effect Size Measurement. Statistical software SPSS version 22 and Excel (2011) were utilized for analysis in this study**.** The Standard Mean Difference (SMD) was selected to calculate the mean effect size for this evaluation. SMD has been calculated by using all available data points in both baseline and treatment or alternatively using only the last three baseline and the last three treatment points (SMD_3, Olive & Smith, 2005). SMD_3 has several limitations including, (a) calculation increases the likelihood of an inflated effect size (Olive & Franco 2008); (b) it may not capture trends within the phases that could impact the mean value and (c) may not provide representation of a treatment effect (Beretvas & Chung, 2008, Rakap, 2015). However, there is no consensus on what type of effect size calculation should be used (Brossart et al., 2006; Parker & Brossart, 2003). As a result, SMD_3 was utilized in this study because of the greater likelihood that a minimum of three data points per phase would be obtained from the file review data.

The formula for calculating this metric is the mean of the last three data points of baseline minus the last three data points from treatment divided by the standard deviation of the baseline mean. Effect size was computed for each participant separately, followed by an overall effect size will be computed for the group. To minimize the inflation of the effect size due to small samples, a correct *d*-statistics (d=1-[3/4N-9] x d) was utilized. In addition, the ESCI-JPP (Finch & Cumming, 2009) excel worksheet was utilized to calculate the confidence intervals of the effect size. Finally, there was 100% agreement between the principal student investigator and student from the companion thesis when entering the data into SPSS and utilized in the statistical analysis in this evaluation.

Results

The results of the study are presented by research question and where necessary, additional post hoc analyses are included.

Research Question #1: What are the presenting problems of the sample targeted during intervention?

The following section will highlight the results of the presenting problem behaviours identified for the entire sample and in each group (a) plans with measurement focused on problem behaviour (b) plans with measurement focused on replacement behaviour.

Presenting Problems. All the clients in this sample had an identified problem behaviour that was targeted for intervention. The frequencies, percentages, and patterns of categories of problem behaviours targeted during intervention are depicted in Table 10. A vast majority (67%) of intervention plans in this sample targeted two or more categories of problem behaviour. While 37% of the participant's interventions were focused on reducing one category of problem behaviour. Disruptive (42%) and dangerous behaviours (40%) had the highest representation across the identified problem behaviour categories.

Refusal alone or in combination accounted for 18% of reported problem behaviour. Finally inappropriate communication on its own or in combination with other problem behaviour comprised 15% of the problem behaviour categories.

Table 8

Frequencies and Percentages of Problem Behaviours of Participants in Overall Sample (n=112)

Problem Behaviours	Frequency	Percentage
Miscellaneous	20	18%
Inappropriate Communication (IC)	8	7%
Disruptive Behaviour	6	5%
Dangerous Behaviours	5	4%
Refusal	3	3%

Combined Problem Behaviours		
Dangerous plus other Problem	19	17%
Dangerous/Disruptive	16	14%
Disruptive plus Refusal	6	5%
Disruptive plus IC	6	5%
Dangerous/Disruptive/Refusal	6	5%
Disruptive plus Miscellaneous	5	4%
Refusal/Disruptive/Miscellaneous	4	4%
Refusal plus Miscellaneous	3	3%
Disruptive/IC/Miscellaneous	3	3%

Target Problem Behaviours and Outcome Measurement. The behaviours that were selected for outcome measurement were extracted from the behaviour intervention plans on file for each participant. As described earlier, behavioural intervention plans were all expected to decrease problem behaviour, and increase a functional replacement skills, however for the vast majority of cases (91 %) the data collected was focused on either the replacement skill or the problem behaviour, with only 9% (n=10) of cases reporting data on both. The remaining analyses in this thesis will present data of the participants divided into 2 groups. One group will include those with measurement focused on the problem behaviour (48% of the plans), while the other group will include those with measurement focused on the replacement behaviour (52% of the plans). The 10 cases with data for both categories were only added to the group with measurement focused on decreasing problem behaviour to avoid overlap.

Plans with measurement focused on the problem behaviour. The problem behaviours targeted in the intervention plans with measurement focused on problem behaviour are summarized in Table 11. A further review of the problem behaviours targeted reveals that 55% of plans included one behaviour category (e.g. dangerous behaviours) and 48% included more

than one behaviour category (e.g. dangerous/Disruptive). Further review of both the single and combined problem behaviours in the sample revealed that both dangerous behaviours and disruptive behaviour were most often targeted in plans with 44% and 38% respectively.

Table 9

Frequencies and Percentages of Problem Behaviours Targeted in Plans Measuring Problem Behaviour (n=54)

Problem Behaviours	**Frequency**	**Percentage**
Miscellaneous	14	26%
Inappropriate Communication (IC)	7	13%
Dangerous Behaviours	4	7%
Disruptive Behaviour	3	6%
Refusal	2	4%
Combined Problem Behaviours		
Dangerous/Disruptive	8	15%
Dangerous plus other Problem Behaviour	7	13%
Dangerous/Disruptive/Refusal	6	5%
Disruptive plus Refusal	1	2%
Disruptive plus IC	1	2%
Disruptive plus Miscellaneous	1	2%
Refusal plus Miscellaneous	1	2%
Refusal/Disruptive/Miscellaneous	1	2%
Disruptive/IC/Miscellaneous	1	2%

Plans with measurement focused on replacement behaviours. The target behaviours for those with intervention plans focused on measurement of the acquisition of replacement skills are summarized in Table 10. In these cases specific data was not collected by the mediators on problem behaviours, as the goals for these plans were to increase replacement behaviours in the absence of the identified problem behaviour. The problem behaviours targeted in these plans were more combined category of dangerous/disruptive had a high representation in the behavioural interventions to increase replacement behaviours. However, dangerous plus other

problem behaviours represented the majority of problem behaviours (26%) plans monitoring skill acquisition, compared to 13% in plans monitoring problem behaviour

Table 10

Frequencies and Percentages of Problem Behaviours Targeted in Plans Measuring Replacement Behaviour (n=58)

Problem Behaviours	**Frequency**	**Percentage**
Miscellaneous	6	10%
Disruptive Behaviour	2	3%
Dangerous Behaviours	1	2%
Inappropriate Communication (IC)	1	2%
Refusal	1	2%
Combined Problem Behaviours		
Dangerous plus other Problem Behaviour	15	26%
Dangerous/Disruptive	8	14%
Dangerous/Disruptive/Refusal	5	9%
Disruptive plus IC	5	9%
Disruptive plus Miscellaneous	4	7%
Disruptive plus Refusal	3	5%
Refusal/Disruptive/Miscellaneous	3	5%
Refusal plus Miscellaneous	2	3%
Disruptive/IC/Miscellaneous	2	3%

Research Question 2: To what extent were service activities and process consistent with the expectations outlined by professional, organizational, and governmental standards in this sample?

In the section below, the ABAS activities and process will be examined to determine the consistency with expectations outlined by MCYS for ABA-based services and organizational/professional standards. Service activities that will be examined include (a)

services provided (e.g. service recipients, duration); (b) assessment (c) treatment (d) service delivery.

Demographic Information

Children and youth with ASD. The sample of 112 were identified for inclusion into this study through a retrospective file review of closed cases (see screening details below) The sample included children and adolescents ranging from 4 to 18 years of age (*M*=8.5, *SD*=3.34) diagnosed with an ASD. The frequencies of children and youth across different age bands are depicted in Figure 1.

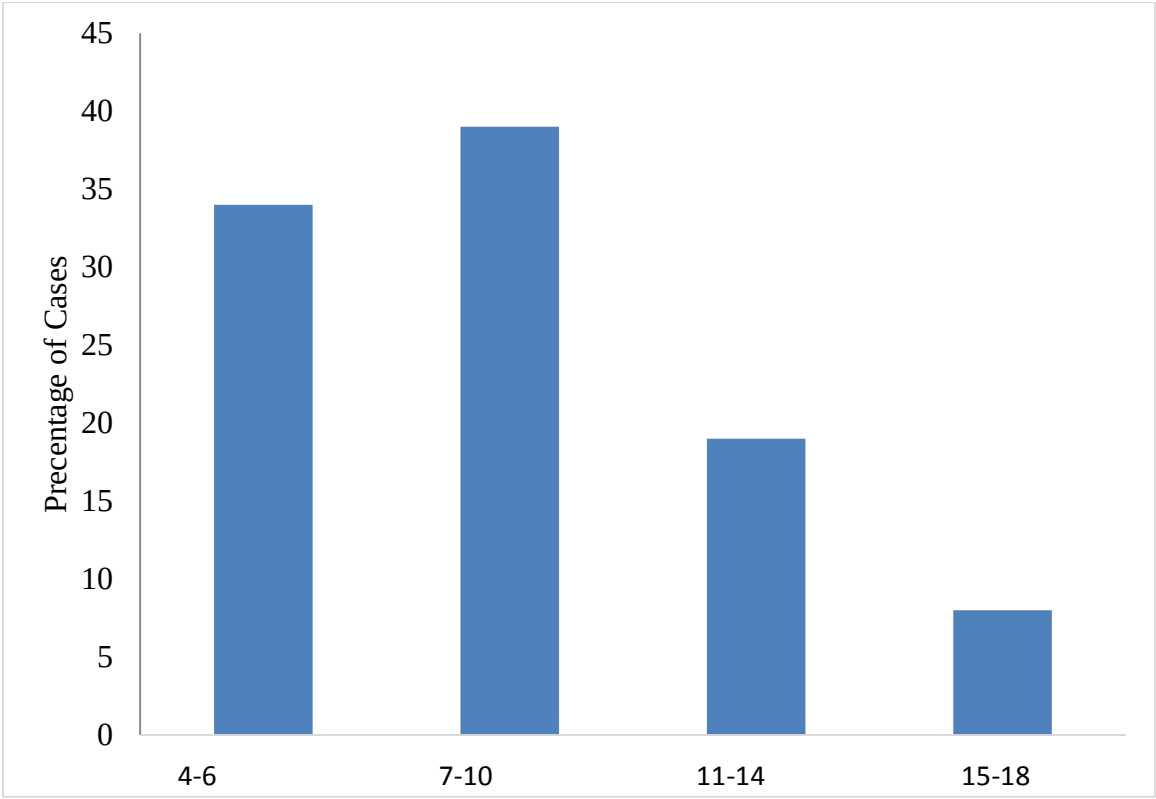

Figure 1. Percentage of sample within age ranges (years) at onset of service (n=112).

The sample included mostly (85.7%) boys and much fewer girls (14.3%). The length of service ranges from or 6 to 20.6 weeks (*M*=11, *SD*=2.95). The majority of the sample (71.4 %)

were receiving services from ABAS for the first time, while the remaining cases included 24.1% receiving services for the second time, and 4.5% receiving services for the third time. On average the number of times receiving services was 1.54 (range=1 to 3, *SD*=0.67).

Caregivers. A total of 163 mediators were trained in the course of service delivery to the 112 children and youth whose files were reviewed in this study. As shown in Table 11 below, the vast majority (89%) of caregivers trained to carry out behavioural intervention plans were biological, adoptive, or foster, parents, and partners of these parents. Please note "parent" in the chart below includes biological, adoptive, foster parents and their partners. For more than half of cases (54.9%) one parent was trained, while both parents were trained in 28.3% of cases. Mothers trained during service were 68% of the total mediators trained. Other mediators included school personnel, grandparents, siblings, and residential staff.

Table 11

Frequencies and Percentages of Caregivers Trained per Participant (n=112)

Type of Caregivers	Frequency	Percentage
One parent	62	55%
Two parents	32	28%
One parent and school personal	5	4%
Two parents and school personal	2	2%
One parent and grandparent	1	1%
Grandparent	2	2%
One parent and sibling	2	2%
Two parents and Sibling	1	1%
Community personal and one parent	1	1%
Group home Staff	2	2%
School Personal	1	1%
Two parents and grandparent	1	1%

Behaviour consultants. The behaviour consultants, whose work was reviewed in the context of this study, had a minimum of an undergraduate degree in psychology and related experience or equivalent with training and expertise in ASD (and/or related disorders) and ABA.

The retrospective design did not permit seeking demographics for each consultant as some had left the organization. A total of 13 consultants developed and implemented behaviour intervention plans included in this file review.

Settings. An objective of ABAS is to enhance caregivers' skills in natural settings in an attempt to reduce reliance on clinical services. Based on Table 12 below, treatment was provided at the participants' home in 95% of the total sample of settings for interventions.

Table 12

Frequencies and Percentages of Settings for Intervention (n=112)

Settings	Frequency	Percentage
Home	99	88%
Home and school	7	6%
Group home	2	2%
Home and community	1	1%
School	1	1%
Daycare program	1	1%
Community	1	1%

Service Recipients. As indicated in the introduction, several requirements outlined by MCYS for ABA-based services includes the provision of services for (1) children diagnosed with ASD up to the age of 18 years (2) provide between 2-6 months of service (3) inclusion of parents/caregivers (e.g. practise the techniques learned, apply the techniques with their children or youth and receive feedback from staff) (4) services provided in a variety of settings (e.g. home, school). Results from this evaluation indicate that overall the ABAS met the expectations for the mandated population.

More specifically, children and youth who participated in services ranged from 4 to 18 years of age (M=8.5, SD=3.34) and were provided services between 6 to 20.6 weeks (M=11,

SD=2.95). Although the majority of cases (99%) fell within the age range outlined by the mandate, one client entered service at the 18 years old. In addition, one client was provided with service for 6 weeks, which is lower than the service duration requirement of 2-6 months. The reasons for early discharge could not be discerned by the file review but the goal for that participant was stated as (met/unmet), Service duration for all the other clients (99%) occurred within the 2-6 months. A total of 163 caregivers were involved and trained during service delivery for 112 participants, thereby meeting the requirement of the caregiver involvement. As reported above, parents represented 89% of the caregivers trained, and 8% of cases included more than one type of caregivers trained (e.g. parent and school personal). Finally, the number of clients where service was provided in more than one setting represented a low percentage (7%) of the settings for service delivery of the total sample (n=112). However, service was provided in familiar settings such as home (95%) and (12%) other community based settings. Overall, the ABAS was in 99% accordance with client and service duration expectations. Caregivers were involved and or trained in 100% of all behavioural intervention plans. Although, a low percentage of cases where more than one setting was utilized for training was reported, familiar settings were primarily utilized during service delivery, therefore, the ABAS was able to meet the final requirement.

Functional Assessment Methodology. Functional behavioural assessments are essential to the development of treatments for problem behaviour (BACB, 2016). Table 13 below illustrates the frequencies and percentages of the FBA methods utilized by the consultants in the ABAS. The most commonly used interviews were the behavioural contextual interview (98%) and the Bio-Psycho-Social interview (98%), which were developed in house and utilized irrespective of the target of intervention. Overall, questionnaires were utilized in 46% of the

sample. More specifically, 41% used one questionnaire, and 27% utilized two questionnaires during the FBA. Further, direct measures were utilized by consultants in 87% of the sample. More specifically, the distribution of the number of direct measures utilized per participant in the sample include: (i) one direct measure (20%) (ii) two direct measures (40%) (iii) three direct measures (27%) (iv) four direct measures (1%). Experimental functional analysis was utilized in 1% of the FBA's in this sample.

Table 13

Frequencies and Percentages of Functional Behaviour Methods Utilized (n=112)

Methods	**Frequency**	**Percentage**
Interviews		
Behavioural Contextual Interview	110	98%
Bio-Psycho-Social Interview	110	98%
Questionnaires		
Questions about Behavioural Function	46	41%
FAST	4	4%
Other:	2	2%
Direct Measures		
ABC Recording by Consultant	90	80%
ABC Recording by caregiver/mediator	68	61%
Informal in-vivo probe	38	34%
ABC Recording by unspecified	6	5%
Functional Assessment Observation Form	1	1%
Experimental Measures		
Structional Analysis	1	1%
Functional Analysis	0	0%

Hypothesized Function. A functional hypothesis is the culmination of the functional behavioural assessment and is fundamental to the design of the behavioural intervention plan for decreasing problem behaviour and increasing replacement behaviour in ABAS. The frequencies and percentages of the functional hypotheses in this sample are depicted in Table 14 below

provides an overview of the frequencies and percentages of the common functional hypothesis that were derived by the consultants. A primary function was identified in the vast majority of cases (91%) and a secondary function was derived in just over 35% of cases. In addition, 94% of the primary functions utilized in intervention plans were one identified function.

Table 14

Frequencies and Percentages of Functional Hypothesis Utilized for Intervention (n=112)

Functions	**Frequency**	**Percentage**
Primary function	102	92%
Secondary function	39	35%
No function identified	10	8%
Primary Function (n=102)		
Escape/Avoidance	38	37%
Tangible	34	33%
Sensory/Automatic	14	14%
Attention	10	10%
Undifferentiated	6	6%
Secondary Function (n=39)		
Attention	21	19%
Tangible	8	7%
Escape	8	7%
Sensory/Automatic	2	2%

Assessment Standards. Overall, the consultants demonstrated a high percentage of agreement (86%-100%) across as outlined in organization and professional standards for assessment. More specifically, every client was referred for treatment of problem behaviour was defined and a functional behavioural assessment was completed. The majority of the sample (86%) utilized two methods (minimum of one indirect and one direct method), whereas 14% utilized one or more indirect assessments with few cases (4%) using only one indirect method.

Finally, the extraction of files, revealed that a functional hypothesis was included in the vast majority of cases (91%).

Measurement Methods. The frequencies and percentages of measurements utilized are reported in Table 15 below. Based on these results, frequency (38%) and percentage (37%) were the most frequently utilized measurements to track intervention progress. Percentage of opportunities represented the vast majority of measurement utilized under the percentage category (92%). Due to the small sample of cases in which latency was used as the measurement (n=3), the time category was utilized to combine both duration and latency data. Duration (minutes) was the most utilized measurement (54%) in the time category. In addition, all scores for the time-based data were converted to minutes to promote consistency in the calculations. Finally, intensity was measured utilizing a scale from 0 (low intensity) to 4 (high intensity).

Table 15

Frequencies and Percentages of Types of Measurements Used (n=112)

Measurement	Frequency	Percentage
Frequency	**43**	**38%**
Percentage	**41**	**37%**
Percentage (Occurrences)	38	
Percentage (unspecific)	2	
Percentage (Partial Interval)	1	
Time	**24**	**21%**
Duration (minutes)	13	
Duration (Seconds)	7	
Latency (minutes)	2	
Latency (seconds)	1	
Duration (hours)	1	
Intensity	**4**	**4%**
Total	**112**	**100**

Treatment Components. The behavioural treatment plans put into place by the consultants consisted of multiple components. The frequencies of the strategies recommended under the

different component categories as well as the overall number of recommendations per person are described in Table 16. Overall the treatment plans had between 4 and 22 recommendations (*M*=9.27, *SD*=3.15) suggesting considerable effort required for training and implementing these plans. Given the agency expectation that behavioural intervention plans will include prevention (e.g. non-contingent reinforcement, alter transitions), skill promotion (e.g. functional communication, differential reinforcement), and behaviour management (e.g. redirection, extinction, reinforce attenuation), it was hypothesized that all cases would have behavioural treatment plans with all three of these components. Though reliability for coding the specific strategies (i.e. redirection) was poor (Awan, 2016), reliability at the categorical level (e.g. prevention) was high and therefore the categorical data are presented here. The vast majority of behavioural intervention plans (91%) included the three treatment components. Further, intervention packages with antecedent and skill building accounted for 8% of the sample, while intervention and skill building only accounted for 1% of the sample.

Table 16

Frequencies and Percentages of Treatment Components Utilized in Intervention (n=112)

Components	Frequency	Percentage
Prevention	434	42%
Skill Building	393	38%
Intervention	209	20%
Total:	1046	

Overall, ABAS had a high percentage of agreement for the treatment standards outlined by the organization and professional standards A vast majority of behaviour intervention plans (91%) included all three types of components (i) Prevention (ii) Skill Building (iii) Intervention.

Replacement Behaviours. As described earlier, an outcome of a functional behavioural assessment is the identification and inclusion of a functionally equivalent or replacement behaviour to increase in an intervention plan. Once a functional behavioural assessment has been completed and a functional hypothesis has been derived, functionally-matched skills were selected for teaching. The frequencies and percentages of identified replacement behaviours are displayed in Table 17 below. Targeting a replacement behaviour was reported in 81% of the entire sample (n=112), including the full sample with data monitoring replacement skill. As indicated below, both other adaptive behaviours and cooperation and were the more prevalent replacement behaviours to increase.

Furthermore, skills assessments were utilized to determine pre-requisite skills and preferences required to teach the replacement behaviours. The frequency of use of skills assessments is similarly elevated for those assessments required by the agency including the functional communication record (99%). Finally, indirect preference assessments were utilized by most consultants (98%).

Table 17

Frequencies and Percentages of Replacement Behaviours across all programs (n=91)

Replacement Behaviours	Frequency	Percentage
Other Adaptive Behaviour	47	43%
Cooperation	24	20%
Tolerance/Waiting	20	18%

Plans with measurement focused on problem behaviours. Table 18 below reports the sample of cases (n=54) whereby the goal and subsequent measurement was to decrease problem behaviours during service. In addition, the table demonstrates the measurement types utilized during treatment. The use of frequency data collection is the most prevalent (65%) across all the

problem behaviours with the exception of disruptive behaviour where time is the most prevalent measurement utilized. The least utilized measurement is intensity, representing 7% of the sample of problem behaviours to decrease. Finally, the mean weeks of service was 11.03 (7.14 to 20.57) for participants in this group.

Table 18

Frequencies and Percentages Measurement Types Utilized Across Plans Measuring Problem Behaviours (n=54)

Problem Behaviours	Frequency	Percentage
Dangerous		
Frequency	4	100%
Dangerous/problem behaviour		
Frequency	7	88%
Intensity	1	13%
Inappropriate Communication		
Frequency	7	88%
Intensity	1	13%
Refusal		
Frequency	3	100%
Disruptive (alone/combine)		
Time	4	45%
Frequency	3	33%
Percentage	2	22%
Miscellaneous		
Frequency	9	64%
Percentage	3	21%
Time	2	14%
Dangerous/Disruptive		
Frequency	5	62%
Intensity	2	25%
Time	1	13%

Plans with measurement focused on replacement behaviours. Due to the fewer categories of replacement behaviours, Table 19 reports both the replacement behaviours targeted and the measurements in this group. As shown below, the goal and subsequent measurement of interventions of 58 cases was to increase replacement behaviours in this during service. Similarly

to the replacement behaviours reported for the overall sample (112), both other adaptive behaviour (54%) and cooperation (27%) are highly represented in this sample. In comparison with the measurement utilized for the plans with measuring problem behaviours frequency is the least utilized (10%), followed by time (29%) and percentage with the highest representation (60%) measurement method across the replacement behaviours. In addition, the mean duration of service for participants in this group was 10.96 weeks (6.00 to 17.57).

Table 19

Frequencies and Percentages Measurement Types Utilized in Plans Measuring Replacement Behaviours (n=58)

Replacement Behaviours	Frequency	Percentage
Cooperation (n=16)		
Percentage	13	81%
Time	2	13%
Frequency	1	6%
Tolerance/Waiting (n=13)		
Time	10	77%
Percentage	2	15%
Frequency	1	8%
Other Adaptive Behaviour (n=29)		
Percentage	20	69%
Time	5	17%
Frequency	4	14%

Service Delivery. The service delivery standards were developed to ensure that services were consistent with best practices and that a high degree of implementation integrity is observed across the consultants. The service delivery standard stipulates the use of a BST model or components of this model when implementing behaviour intervention plans with mediators. However, a low percentage (16%) (n=18) of consultants reported the utilization of the BST model by name, but, higher percentages of consultants reported using BST components: (i)

instruction (95%) (n= 106); (ii) live modelling (22%) (n=25); (iii) role-playing (29%) (n=32); (iv) feedback (20%) (n=22).

Overall, the service activities and process of ABAS in this sample are consistent with the expectations of government, organizational and professional standards. As highlighted in this section, high percentages were observed in the service recipients, assessment and treatment activities of ABAS. The results for the service delivery standards were lower, however, the instruction component yielded a high percentage.

Research Question #3: Is there meaningful improvement of presenting problems at the conclusion of services for this sample?

Service outcomes were measured by percentage of behaviour change and effect size calculations for both groups (a) plans with measurement focused on problem behaviour and (b) plans with measurement focused on replacement behaviour.

Behaviour Percentage Change. The use of behaviour percentage change has been utilized in studies to display changes to behaviour, comparing treatment mean to baseline means for both problem and replacement behaviour (Feldman et al., 2002; Carr et al., 1999). For problem behaviour percentage of behaviour change was calculated by subtracting the mean of the last three treatment points from the mean of the last three baseline points then dividing by the mean of last three baseline points (Carr et al. 1999). Percentage of behaviour change for changes in replacement behaviours was calculated by subtracting the mean of the last three baseline data points from the mean of last three treatment data points then dividing by the mean of the last three baseline data points (Feldman et al., 2002). Please note, for both groups a 100% cut-off was utilized per case. More specifically, cases with percentages above 100% + or – were provided a score of a 100% with the applicable sign. Table 20 below highlights positive percentage

behaviour change occurring for plans measuring replacement behaviours (84%) and plans measuring problem behaviour (57%). Using Table 21 to summarize, moderate improvement was observed for plans with measurement focused on problem behaviours (44%-77%) compared to substantial improvement of plans with measurement focused on replacement behaviours (77%-100%). In addition, problem behaviours that include (a) dangerous/other problem behaviour (b) inappropriate communication (c) refusal reported higher percentage changes (73%-78%) compared to the overall mean of the group. Further, the ranges indicate these three behaviour categories were also only categories to report that problem behaviour did not increase or get worse after intervention. In comparison, 2 or the 3 replacement behaviours targeted in plans measuring replacement behaviour, included a participant where the skill decreased or get worse after intervention. Further, cooperation was the only replacement behaviour with no participant reporting a decrease of the replacement behaviour.

Table 20

Percentage of Behaviour Change from baseline to treatment of Problem Behaviours (n=54) and on Replacement Behaviours (n=58)

Problem Behaviours	**N**	**Mean**	**SD**	**Minim**	**Maxim**
Dangerous (alone)	4	43.68	70.33	-45.45	100
Dangerous/problem behaviour	8	77.08	35.83	0	100
Inappropriate Communication	7	75.91	13.84	63.16	100
Miscellaneous	14	45.48	69.03	-100	100
Dangerous/Disruptive	8	55.72	69.68	-100	100
Disruptive (alone/combination)	10	47.62	58.21	-100	100
Refusal (alone/combination)	3	73.21	23.35	57.14	100
Plans Measuring Problem Behaviour	**54**	**57.43**	**55.37**	**-100.00**	**100.00**
Replacement Behaviours					
Cooperation	16	79.64	40.17	-31.82	100.00

Tolerance/Waiting	13	100.0	.00	100.00	100.00
Other Adaptive Behaviour	29	79.44	41.25	-57.14	100.00
Plans Measuring Replacement Behaviour	**58**	**84.12**	**36.53**	**-57.14**	**100.00**

As highlighted in Table 22, further analysis of plans with measurement focused on problem behaviour reveals that 68% of participants experienced a 75% or greater reduction or moderate improvement in problem behaviour. Of those, 30% of participants experienced a 100% reduction in problem behaviour. Finally, 31% of the participants experienced a reduction of behaviour of 50% or less. In comparison, 88% of participants experienced an increase of 100% in replacement behaviour or substantial improvement as part of plans with measurement focused on replacement behaviour.

Table 21

Level of Improvement from Baseline across Percent Change Ranges

Percent	*Level of Improvement*
<0	Worse
0% - 25%	Minimum Improvement
26% - 50%	Moderate Improvement
51% - 75%	Marked Improvement
76% - 100%	Substantial Improvement

Note: Scale adopted from the global assessment of improvement from baseline scale (Dhuin, 2009)

Table 22

Percentages of Percent Change Intervals for Both Plans (n=112)

Percent Change	**Frequency**	**Percentage**
Plans Measuring Problem Behaviour (n=54**)**		
<0	5	9%

0% - 25%	6	11%
26% - 50%	6	11%
51% - 75%	12	22%
76% - 100%	25	46%
Plans Measuring Replacement Behaviour (n=58)		
<0	3	5%
0% - 25%	2	3%
26% - 50%	1	2%
51% - 75%	1	2%
76% - 100%	51	88%

Effect Size Calculations

As previously described, behaviour intervention plans used in the ABAS are expected to reduce problem behaviours, though this may not be explicitly measured for all participants. The SMART goal determines the measurement and direction of the behavioural intervention plans. As previously described, the participants were placed in one of two groups (a) Plans with measurement focused on replacement behaviour (n=58) and (b) Plans with measurement focused on problem behaviour (n=54). In order to ensure that effects could be detected, within these groups, sub-groups were created based on the target behaviour categories found in each group (e.g. Increase group: cooperation). As described in this evaluation, an effect size was first calculated for each participant, followed by overall effect sizes for both groups and their corresponding sub-groups. The effect sizes for these subgroup will be interpreted using Cohen's *d* (1998) levels *d*=0.2 small, *d*=0.5 medium and *d*=0.8 large. Table 23 displays the mean effect size calculations for both the increase and decrease behaviour interventions across the sub-groups. However, the effect sizes of a total of 27 participants were not calculated due to standard deviations of zero. Behaviour percent change was computed for this group and will be included in the results below. Further, the table displays the adjusted *d*-statistic for small sample size and

confidence intervals. Given the small sample size for the refusal, an adjusted *d*-statistic was not calculated for that sub-group.

Table 23

Mean Effect Size for Both Plans Across Sub-Groups (n=85)

Problem Behaviours	**N**	***D***	**SD**	**Min**	**Max**	***d****	**CI**	
Dangerous	4	.67	1.03	-.61	1.66	.384	-0.97	2.31
Dangerous/problem behaviours	8	1.98	1.27	.00	3.67	1.72	0.92	3.04
Dangerous/Disruptive	6	1.79	1.02	.48	2.89	1.43	0.72	2.86
Disruptive	9	.36	5.07	-13.75	2.88	.323	-3.53	4.26
Refusal	3	2.11	1.68	.87	4.04	2.11	-2.06	.628
Inappropriate Communication	7	2.55	1.19	.95	4.04	2.14	1.44	3.65
Miscellaneous	13	.97	2.93	-5.46	6.40	.93	-0.80	2.74
Plans Measuring Problem Behaviour	**50**	**1.25**	**2.81**	**- 13.75**	**6.40**	**1.23**	**0.45**	**2.05**
Replacement Behaviours								
Cooperation	11	-2.79	2.94	-7.51	.93	-2.55	-4.76	-0.81
Tolerance/Waiting	7	-20.23	25.68	-77.36	-1.15	-17.04	-29.05	-11.41
Other Adaptive Behaviour	17	-4.20	6.69	-28.29	-2.31	-3.98	-7.59	-0.80
Plans Measuring Replacement	**35**	**-6.96**	**13.63**	**-77.36**	**2.31**	**-6.80**	**-11.64**	**-2.27**

Note: *d***=** effect size**;** *d****=**adjustment effect size, CI= Confidence Intervals (p<.05)

Plans with measurement focused on problem behaviour. Overall decreasing interventions had large effects (Cohen's $d > 0.8$) and were statistically significant based on Confidence Intervals (CI) (the 95% CI's did not contain zero). Similarly, 5 of the 7 sub-groups had large effects based on the adjusted *d* statistic. However, only (a) dangerous/other problem behaviour (b) Dangerous/Disruptive (c) Inappropriate communication were statistically significant (the 95% CI's did not contain zero).

Plans with measurement focused on replacement behaviour. Overall increasing behavioural interventions had large effects (Cohen's $d > 0.8$) and were statistically significant based on Confidence Intervals (CI). Similarly, all the sub groups reported a large effect size ($d > 0.8$) and were statistically significant (the 95% CI's did not contain zero). As a result, a (-)

accompanying an effect size is preferred as this suggests an increase was observed from baseline to treatment.

Participants Not Included. Twenty-seven participant's effect size computations yielded errors and were not included in the results above due to standard deviations of zero. Instead, percent change calculation was utilized to examine the outcomes of these participants. Four out of the 27 participants had measurement focused on problem behaviour. Two of these participants had low overall behaviour percent reduction (31%) with ranges of -100% to 100%, while 2 of these participants experienced a 100% reduction**.** In addition, for 23 of the 27 participants, replacement skill acquisition was measured. This group experienced a mean increase in replacement behaviours of 93.47% with ranges of 0% to 100%. Further, 91% of this sample had a 100% increase in replacement behaviours.

Discussion

The purpose of this retrospective study was to examine the clinical processes used and the behavioural outcomes of children and youth with ASD who participated in the ABAS. This was accomplished by examining the presenting problem behaviours of the sample, the assessment and intervention plans used to treat problem behaviour, and the behavioural outcomes achieved by the participants.

The following will begin with a discussion based on the research questions followed by a discussion of the strengths, limitations, future directions and conclusions.

What are the presenting problems of the sample targeted during intervention?

As highlighted in the results section, dangerous and disruptive behaviours were the most prevalent problem behaviours in the overall sample (112) and in plans with measurement focused on problem behaviours (54). These results are consistent with the literature on the prevalence of problem behaviours exhibited by children and youth with ASD and that are targeted in behavioural interventions (Horner et. al., 2002; Hattier et al. 2002, National Autism Centre, 2015). Although specific problem behaviours such as throwing items and aggression to others were not measured individually in this study, they are included in the definitions for disruptive and dangerous behaviours respectively. Single categories of problem behaviours (e.g. dangerous behaviours) were represented in both the overall sample and in plans measuring problem behaviours.

Further, provided that many participants exhibited several categories of problem behaviours highlights the importance of an effective assessment and intervention components of service delivery.

To what extent were service activities and process consistent with the expectations outlined by professional, organizational, and governmental standards in this sample?

In the following section, the ABAS activities and process will be highlighted to determine the agreement implementation fidelity (Fettig et al., 2015) with expectations outlined by the MCYS for ABAS; service and professional standards.

MCYS Expectations. The results revealed that the ABAS met several of the expectations of ABA-based Services outlined by the MCYS. Services were provided to the appropriate age group, and within the service duration expectations. Moreover, the vast majority of the caregivers were informed, provided instructions and several trained using other methods (e.g. BST, role-play) to implement the behaviour interventions for their child. As indicated, the majority of services in this sample involved the training of parents in home settings. Although, training occurred in other settings such as school and community based settings, home settings represented the largest proportion of training settings. However, the number of caregivers who are natural agents in the child's life involved in services ensures the continuation and possible maintenance and generalization of the strategies trained. It should be noted that the behaviour intervention plans did include sections and steps for both maintenance and generalization, however, for the purpose of this evaluation this data was not extracted. Overall, services provided are consistent with several of the expectations for ABA-based services outlined by the MCYS.

Service and Professional Expectations. Standards of best practice are essential in the provision of ethical and effective service delivery. The ABAS leadership responsible for the clinical oversight of services (e.g. Management, BCBA's, Psychologists), developed and monitored the congruency of these standards with the service activities and process of ABAS.

The results highlight the high percent agreement and integrity of the consultant performances across all three components. The vast majority of behaviour intervention plans for both increasing and decreasing behaviour included a problem behaviour (100%), completed a FBA utilizing more than one method (86%), and identifying a primary function (91%). The results of prevalence of primary functions targeted in behavioural interventions for this sample is consistent with the functions identified for this population in the literature (Love et al. 2009).

The identification of the function(s) that maintain the problem behaviour and inclusion into behavioural intervention plans is considered best practice and leads to improved behavioural outcomes (BACB 2016; Carr et al., 1999). The results in this evaluation reveal 98% of the consultants used a method as part of the FBA. In addition, approximately half of the sample utilized a questionnaire, whereas, the use of experimental measures was the least commonly used method. These results are consistent with results of surveys of behaviour analyst's utilization of FBA methods in practice (Oliver, Pratt & Normand, 2015). In the survey, most of the respondents reported the use of informant and descriptive assessments more frequently than functional analysis, and majority indicated that they "never" or "almost never" use functional analysis to determine the function for problem behaviour (Oliver, Pratt, & Normand, 2015). A functional analysis is considered to be the gold standard of FBA's, and more effective at determining function compared to both indirect and direct methods (Hanley et. al. 2003; Thompson & Iwata, 2007). However, an implication of this study is that the results derived from the combination of both indirect and descriptive methods can be utilized to develop effective intervention plans and produce positive behavioural outcomes.

High percent of agreement was observed with the expectations of treatment components for increase and decrease behaviour intervention plans. The results indicated that 93% of

programming included all three treatment component parts (e.g. prevention, skill building, & intervention). In addition, the most common components utilized were prevention and skill building components. There was a proactive focus in treatment utilized in this sample. The number of recommendations per program had a wide range with 22 on the high end. Although the average was 9.27 recommendations per program, these results suggest that considerable expectations are placed on the caregiver to implement behaviour intervention plans. Similarly, considerable effort is required by the behaviour consultant to train caregivers the recommendations. Although there is no evidence that the caregivers implemented all the recommendations, the positive behavioural outcomes reported in this study would suggest that caregivers were able to implement strategies (e.g. 1 or more). Consistent with the literature, implications of these results are caregivers when provided with training are able to implement strategies and have a positive impact on the behavioural outcomes (e.g. Suess et al., 2014, Feldman et al., 2002).

These results contribute to the literature on multi-component interventions. Similar to previous effectiveness or "real life" studies (Feldman et al. 2002), the results of this evaluation demonstrate that natural mediators can successfully implement multi component interventions in applied settings. Although, the treatment adherence of the mediator's implementation of intervention was not collected, the behavioural outcomes reported in this evaluation suggests that some components if not all were implemented correctly by the mediator. These results contribute to the literature on mediator implemented multi-component interventions to decrease problem behaviours for children and youth with ASD.

Replacement Behaviour. The results in this evaluation revealed that 81% or 91 cases of the overall intervention plans (112) included a replacement behaviour. Cooperation and other

adaptive behaviours were frequently targeted in both the overall sample, and in the subsample for whom measurement focused on replacement behaviours. Similar to the results outlined in the report by National Autism Center (2015), a focus on teaching adaptive behaviours was common in the replacement skills identified in this study. Cooperation skills were taught in this sample, and were not included explicitly in the National Standards Report (National Autism Centre, 2015). The low reliability on specific replacement skills and strategies (Awan, 2016), limited the ability to provide results on a more detailed level however, these results contribute to the literature on the type of replacement behaviours that are targeted in behaviour interventions in applied settings.

Measurement. Based on the results, caregivers utilized all dimensions of measurement with frequency and percentage representing 75% of measurement utilized in the overall sample.., similarly, frequency was utilized in 65% of measurement systems for problem behaviours, but percentage of occurrence was the measurement system used in 60% of measurement systems for replacement skills. However, what was not indicated in the reports extracted were the reasons for choosing specific measurement method based on the target behaviour. As well, the integrity of the implementing the type of measurement was not indicated in the files reviewed. However, these results demonstrate that in applied settings, caregivers were able to collect data across several types of measurements to track intervention progress.

Similarly, the vast majority of consultants (95%) met the service delivery expectations of providing instructions and explaining the intervention to caregivers. In contrast, results for the utilization of the BST model and other methods (e.g. role-play) was reported less (range 20%-29%). However, it should be noted that at the time of the extraction, the expectation to document the type(s) of methods utilized to train the caregivers was not indicated on the service provision

report. Overall, these results reveal that a high implementation adherence and congruence with the standards of practice outlined by both the MCYS and organization/professional standards and the ABAS and process. In addition, these results highlight the importance of examining the service activities of a program to determine whether expectations are being met. More importantly, as indicated by Posavic (2015), the results of the service activities and process of ABAS provides interested stakeholders such as funding sources with the information to make program level decisions.

Is there meaningful improvement of the presenting problems at the conclusion of services?

Service outcomes were computed by utilizing percentage of behaviour change and effect size computations for both: (a) Plans with measurement focused on problem behaviours (b) Plans with measurement focused on replacement behaviours. Results demonstrate positive percentage change for both groups. More specifically, moderate improvements were observed for plans measuring problem behaviour. The percentages of behaviour reduction for the plans with measurement focused on problem behaviour are similar those reported by Carr et al. (1999). However, Carr et al. (1999) included specific target behaviour behaviours (e.g. aggression, self-injurious behaviour), and ranges of percent reduction of problem behaviour across studies (366) included ranges of 38% to 62%. In addition, Carr et al. (1999) reported single type behaviours (e.g. aggression) had percentage reduction of 55.2%, compared to 44% for combination types of problem behaviour. In comparison, the majority of problem behaviour categories utilized in this study would more likely be considered a “combination type” according to the aforementioned descriptions. Similarly, plans with measurement focused on replacement behaviours produced significant improvements. In comparison, the behaviour percentage change was greater for plans measuring replacement skills then for plans measuring problem behaviours. The implications of

these results are that replacement behaviours compared to problem behaviour are more likely to significantly improve during a service duration of an average of 11 weeks. Similar to Feldman et al. (2002) the results in this evaluation indicate that in both types of intervention plans, caregivers were able to successfully reduce problem behaviour and increase replacement behaviour. Further analysis would be required to determine variables that influence the difference in percent change between increasing a replacement skill compared to decreasing problem behaviour. However, these results demonstrate the positive improvements in behaviour change for participants who were provided ABAS.

Similarly**,** positive outcomes were observed with the effect size calculations for both types of intervention plans. More specifically, both the plans with measurement focused on problem behaviour and plans with measurement focused on replacement behaviour groups had overall large effect sizes ($d > 0.8$) and were statistically significant (CI). The majority of the sub-groups for both increase and decrease programs had large intervention effects when interpreting using the parameters by Cohen's *d*. Unfortunately, a review of the literature reveals that there is dearth of evaluation of programs that utilizes effect size calculations to determine intervention effects for behavioural outcomes for children with ASD. The large effect sizes revealed in the results by both groups and the majority of their corresponding sub-groups demonstrate the improvements observed as a result of participating in ABAS. Moreover, these results confirm the hypothesis that positive outcomes were observed for participants in this sample who received ABAS for help with problem behaviour. Therefore, the outcomes computed by both the percentage change in behaviour and in effect size provides evidence to support the conclusion that the participants in this sample who received ABAS observed positive outcomes by the end of service delivery (average 11 weeks).

Strengths

An objective of this study was to evaluate the behavioural outcomes of the ABAS using the single subject data utilized during treatment. Thus, a considerable strength of this study was the demonstration of a program evaluation with a large sample (n>100), utilizing single subject data to determine behavioural outcomes. This study provides encouraging results that suggest that the majority of participants in this sample who received ABAS using a mediator model observed meaningful gains and positive outcomes in an average of 11 weeks. As previously highlighted, a dearth of literature exists of program evaluations that both utilize single subject data (Burns, 2014) and evaluated services for a large sample children with ASD receiving parent-implemented behaviour intervention. These findings contribute to the literature that parent-implemented behaviour intervention is effective in reducing the problem behaviour exhibited by their child with ASD.

Another strength was the use of a systematic file review tool and that good reliability was obtained for over 40% of the cases in the sample. The BATT, allowed for a thorough file extraction that provided a wealth of information utilized in this evaluation. The structure of the BATT is congruent with commonly used behavioural intervention plans and programs that minimizes the response effort of the abstractors during the file review. This evaluation also provides a model of utilizing clinical data obtained from case files in meaningful way to demonstrate outcomes of interventions. Finally, evidence from this study suggests that community behaviour consultants delivered services according to professional and program requirements, that clients received the expected services, and that their caregivers could be trained to provide effective behavioural interventions to their children or youth with ASD thereby reducing problem behaviour and/or increasing replacement skills.

Limitations

Although, positive implications and outcomes have been highlighted, several limitations should be considered. Methodological limitations exist with the use of retrospective evaluations using clinical data and information (Anthony, 1992). For example what is indicated in the file may or may not be accurate. Although the extraction revealed that none of the functional behaviour assessments included a functional analysis, no explicit instructions were provided on the reports reviewed that would prompt the consultant to document the utilization of this method. Furthermore, whether information is documented or not documented, when utilizing a retrospective method, the accuracy of that information may be questioned. More specifically, with the retrospective analysis, the omission of information (i.e. method utilized) may not reflect the accuracy of service provided.

Another potential limitation is that principal student investigator is a BCBA supervisor and two of the research assistants are behaviour consultants in the ABAS and were not blind of the objectives of this evaluation when screening and abstracting data. However, high reliability with the extraction and coding among the other research assistants, principal student investigator of the companion thesis (Awan, 2016) with the principal student investigator suggests that the information included in this evaluation may be unbiased.

In addition, a limitation may exist in SMD calculation. More specifically, Beretvas & Chung (2008) identified some methodological issues when utilizing a SMD of 3 treatment/baseline data points. First they suggest that SMD_3 may lead to false detection of a treatment effect due to natural development. Second, they note that a positive (or slight) trend in the baseline and or treatment phases would go undetected. However, effect size are themselves unambiguous, and the magnitude of results should be compared to the model used (Parker,

Brossart, Vannest, & Long, 2005). According to Olive & Franco (2008), effect size calculation has the benefits such as being able to be used for increasing and decreasing interventions, not excluding data due to overlapping data, interpretable results comparable with Cohen's *d*, and the calculation is simple. Given that there is no consensus on what type of effect size calculation should be used (Brossart et al., 2006; Parker & Brossart, 2003), for many programs in community based settings, financial resources for evaluation personnel may be limited (Carmen, 2007), utilizing SMD to calculate intervention effect and percent of behaviour change may provide an effective option. As such, results from this evaluation can be compared to other studies that have utilized *d* statistic to calculate effect sizes, which according to (Olive & Franco, 2008) is a strength of used the standardized mean difference.

Another limitation is the interpretation of "replacement or functionally equivalent behaviours" for plans with measurement focused on replacement behaviours in this evaluation. More specifically, based on the SMART goals of these behavioural intervention plans, the objective is an increase of a replacement behaviour, based on the function of the problem behaviour, and in the absence of or a subsequent decrease of problem behaviour. However, in these intervention plans, there was no data collected and or graphed on the problem behaviour to confirm that increases in the targeted skills resulted in a collateral decrease in the problem behaviour. This prohibits the conclusions that the replacement behaviour was functionally equivalent to the target problem behaviour and that there was a problem behaviour reduction for the participants in this group. Although program standards did not require data to be collected on both problem behaviour and functional replacement skills, best practices would include measurement of both (BACB, 2014)

Furthermore, a limitation may exist with the generalization of these results to participants outside this sample within the home agency and in other agencies given the inclusion criteria of files requiring daily data collection. More specifically, in the sample utilized in this study, caregivers collected the daily data collection. The positive outcomes highlighted in this study may have been influenced by the caregiver level of adherence or possible mediator bias observed with data collection. However, this study required the data collection in the files in order to calculate the percentage of behaviour change and effect sizes of interventions.

In addition, a limitation may exist in the lack of mediator interobserver agreement (IOA) on the data utilized in the behaviour percentage change and effect size calculations. IOA is the percentage of intervals in which independent observers agree on the recording of the same event (or dependent variable) (Horner et al., 2005). IOA is utilized in ABA as an indicator of measurement quality (Cooper, Heron & Heward, p. 113, 2007). However, the reports that were utilized in this evaluation did not include a section or prompt for consultants to document IOA data points. Therefore, if IOA was collected, it was not documented on the reports used for this evaluation.

Finally, on average the number of times a participant receiving services was 1.54 (range=1 to 3, *SD*=0.67). A limitation may exist in not conducting a further analysis on the potential impact of the number of times a participant receives services on the impact of the outcomes. Future evaluation should consider conducting a further analysis on the aforementioned variables.

Future Directions

As described by Posavic (2015), an integral component of evaluation is to provide feedback to all stakeholders involved in the program or organization. The following section will highlight feedback for future directions across the different stakeholders associated with the ABAS. Based on this evaluation, feedback for the ABAS include: (a) Increase utilization of functional analysis to increase effectiveness of assessment: Although the results indicate that positive outcomes were achieved highlighted by the large effects in both groups, with the utilization of an indirect and direct assessment methods without functional analysis. The overwhelming evidence does support the use of a functional analysis in determining the function of behaviour (Hanley et al., 2003; Hanley, Iwata & Thompson, 2001, Iwata et al., 1994); However, in 31% of the participants in the plans with measurement focused on problem behaviour experienced a 50% or lower reduction in problem behaviour and 9% experienced an increase in their problem behaviour. The utilization of a functional analysis in the cases listed above may have increased the likelihood to a positive outcome (b) Introducing a mechanism to prompt consultants to document specifics of service delivery such as the training strategies utilized; (c) Embed treatment adherence checks for treatment integrity of program implementation by mediators; (d) Include a process to collect inter-observer agreement of the data collected and graphed per participant; (e) Examine relationships between service duration, outcomes, assessment methods, treatment types and topography of problem behaviour to obtain results that may increase the effectiveness and quality of services. More specifically, an analysis of the behavioural outcomes for specific treatment recommendations, presenting problem behaviours and service delivery duration could provide useful information in streamlining

service delivery to ensure that the ABAS provides flexible and more individualized service to meet the needs of children and their families.

Future studies on program evaluations should include the utilization of the BATT for retrospective and or prospective analysis, to gather information that would allow for the identification of participant and or treatment variables that may impact and or predict behavioural outcomes. Given time and resources constraints for evaluations, agencies and or service should endeavour to develop an infrastructure for practitioners that would allow for simple entry of data points that would later be utilized in a calculation of intervention effects (e.g. an app). The BATT could provide the basis for that infrastructure, and could be customized to meet individual agency expectations beyond best practices.

Future recommendation for government officials is the consideration of utilizing client clinical data when examining the intervention effectiveness of services based on ABA provided to children and youth with ASD. More specifically, clinical data can support decisions around service needs based on client and family profiles. In addition, outcomes for this "effectiveness" evaluations provides decision makers evidence of the impact of intervention in real settings. Similarly, an evaluation that includes a cost-benefit analyses incorporating costs, services activities and outcomes would further provide decision makers information to make informed decisions about the effectiveness of ABAS.

Summary and Conclusion

As highlighted in previous sections, the objectives of an evaluation described by Posavic (2015) is to assess the needs of a group (ii) examine the service activities and processes of the program (iii) examine the outcomes of the program were utilized in this evaluation.

Given the increase of funding for services for children with ASD and fiscal responsibility, funding sources and service providers would be beneficiaries of results that utilize clinical data or data utilized during treatment as part of determining the effectiveness of a service. However, a dearth of published evaluations for services for children and youth with ASD have utilized clinical data to determine the effectiveness of services. Common evaluation practices, exclude the utilization of a tool similar to the BATT that captures information including assessment, treatment features and clinical data that would allow an analysis of variables that impact intervention effects per participant and or as a group. Overall, this evaluation demonstrated that the ABAS met the requirements of provincial expectations; provided service activities to participants consistent with best practices in ABA and achieved overall positive outcomes for children and youth with ASD and their families in this sample.

References

Alter, P. J., Conroy, M. A., Mancil, G. R., & Haydon, T. (2008). A comparison of functional behavioral assessment methods with young children: Descriptive methods and functional analysis. *Journal of Behavioral Education, 17,* 200-219.

Athens, E. & Vollmer, T. (2010). An investigation of differential reinforcement of alternative behavior without extinction. *Journal of Applied Behavior Analysis, 43*(4), 569-589.

Awan, A. (2016). *Assessment of the Inter-Rater Reliability of the Behavioural Assessment and Treatment Taxonomy (BATT)* (Unpublished Masters Thesis). Brock University, St. Catherines, Ontario.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*(1), 91–97.

Baker-Ericzén, M.J., Stahmer, A.C., & Burns, A. (2007). Child demographics associated with outcomes in a community-based pivotal response training program. *Journal of Positive Behavioral Interventions, 9,* 52-60.

Behaviour Analyst Certification Board. (2014). *Professional and ethical compliance code for behaviour analysts.* Retrieved April 7, 2016, from the Behaviour Analyst Certification Board Web site: http://bacb.com/ethics-code/.

Beretvas S. N., Chung H. (2008). A review of meta-analyses of single-subject experimental designs: methodological issues and practice. *Evidence Based Community Assessment Intervention, 2,* 129–141.

Bijou, S. W., Peterson, R. F., & Ault, M. H. (1968). A method to integrate descriptive and experimental field studies at the level of data and empirical concepts. *Journal of Applied Behavior Analysis, 1*, 2.

Burns, C. (2014). Does my program really make a difference? Program evaluation utilizing aggregate single-subject data. *American Journal of Evaluation, 36* (2), 191-203.

Brossart, D. F., Parker, R. I., Olson, E. A., & Mahadevan, L. (2006). The relationship between visual analysis and five statistical analyses in a simple AB single-case research design. *Behavior Modification, 30, 531–563.*

Cameron, M. J., Ainsleigh, S. A., & Bird, F. L. (1992). The acquisition of stimulus control of compliance and participation during an ADL routine. Behavioral Residential Treatment.

Carman, J. G. (2007). Evaluation practice among community-based organizations research into the reality. *American Journal of Evaluation, 28* (1), 60-75.

Carr, E. G., Durand, V. M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis, 18 (*2), 111–126.

Carr, E. G. (1994). Emerging themes in the functional analysis of problem behavior. Journal of *Applied Behavior Analysis*, *27*(2), 393-399.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Condillac, R. A. (2009a). Evaluating processes and outcomes for behaviour therapy services: Phase 1 - Developing the Measures. Research Proposal Submitted to Brock University Research Ethics Board, St. Catharines, ON: Brock University.

Condillac, R.A. (2009b). Impact on Quality of Life (IQOL). Unpublished Measure. St. Catharines, ON: Brock University.

Condillac, R.A. (2009c). The Target Behaviour Severity Scale (TBSS). Unpublished Measure. St. Catharines, ON: Brock University.

Condillac, R.A. (2009d). The Behavioural Assessment and Treatment Taxonomy (BATT). Unpublished Measure. St. Catharines, ON: Brock University.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis*. Columbus, OH: Merrill.

Dhuin, J. C. (2011). *Therapy regimen for treating acne related diseases*. Retrieved from http://www.google.com/patents/EP2288349A1?cl=en.

DiGennaro, F. D., Martens, B. K., & Kleinmann, A. E. (2007). A Comparison of Performance Feedback Procedures on Teachers' Treatment Implementation Integrity and Students' Inappropriate Behavior in Special Education Classrooms. *Journal of Applied Behavior Analysis, 40*(3), 447–461.

Dyer, K., Martino, G., & Parvenski, T. (2006). The River Street Autism Program: A case study of a regional service center behavioral intervention program. *Behavior Modification, 30*, 925-943.

Ducharme, J. M., & Feldman, M. A. (1992). Comparison of staff training strategies to promote generalized teaching skills. *Journal of Applied Behavior Analysis, 25*(1), 165–179. http://doi.org/10.1901/jaba.1992.25-165

Durand, V. M., & Carr, E. G. (1991). Functional communication training to reduce challenging behavior: Maintenance and application in new settings. *Journal of Applied Behavior Analysis, 24*(2), 251-264.

Eisenhower, A. S., Baker, B. L., & Blacher, J. (2005). Preschool Children with Intellectual Disability: Syndrome Specificity, Behaviour Problems, and Maternal Well-Being. *Journal of Intellectual Disability Research : JIDR, 49* (9), 657–671.

Feldman, M. A., Condillac, R. A., Tough, S. E., Hunt, S., & Griffiths, D. (2002). Effectiveness of community positive behavioral intervention for persons with developmental disabilities and severe behavioral challenges. *Behavior Therapy, 33,* 377-398.

Fettig, A., Ostrosky, M., M. (2011). Collaborating with parents in reducing children's challenging behaviors: Linking functional assessment to intervention. *Child Development Research*, 2011 (2011).

Fettig, A., Barton, E. E. (2014). Parent Implementation of Function-Based Intervention to Reduce Children's Challenging Behavior: A Literature Review. *Topics in Early Childhood Special Education, 34 (*1), 49-61.

Fettig, A., Schultz, T. R., Sreckovic, M. A. (2015). Effects of coaching on the implementation of functional assessment-based parent intervention in reducing challenging behaviors. *Journal of Positive Behavior Interventions, 17,* 3.

Fisher, Wayne W., David E. Kuhn, and Rachel H. Thompson. (1998). Establishing discriminative control of responding using functional and alternative reinforcers during functional communication training. *Journal of Applied Behavior Analysis, 3*(4), 543-560.

Fiske, K. E. (2008). Treatment Integrity of School-Based Behavior Analytic Interventions: *A Review of the Research. Behavior Analysis in Practice, 1*(2), 19–25.

Griffiths, D. M., & Gardner, W. I. (2002). The integrated biopsychosocial approach to challenging behaviours. In D. M. Griffiths, C. Stavrakaki, & J. Summers (Eds.), Dual

diagnosis: An introduction to the mental health needs of persons with developmental disabilities (pp. 81–114). Sudbury, ON: Habilitative Mental Health Resource Network

Gillett, J. N., & LeBlanc, L. A. (2007). Parent-implemented natural language paradigm to increase language and play in children with autism. *Research in Autism Spectrum Disorders, 1*(3), 247-255.

Hanley, G. P., Iwata, B. A., & Thompson, R. H. (2001). Reinforcement schedule thinning following treatment with functional communication training. *Journal of Applied Behavior Analysis, 34*(1), 17-38.

Hagopian, L. P., Fisher, W. W., & Legacy, S. M. (1994). Schedule effects of non-contingent reinforcement on attention-maintained destructive behavior in identical quadruplets. *Journal of Applied Behavior Analysis, 27*(2), 317-325.

Hanley, G. P., Iwata, B. A., & McCord, B. E. (2003). Functional analysis of problem behavior: a review. *Journal of Applied Behavior Analysis, 36* (2), 147–185.

Hartley, S. L., Sikora, D. M., & McCoy, R. (2008). Prevalence and risk factors of maladaptive behaviour in young children with Autistic Disorder. *Journal of Intellectual Disability Research : JIDR, 52 (*10), 819–829.

Hattier M. A., Matson J. L., Belva B. C., Horovitz M. (2011). The occurrence of challenging behaviours in children with autism spectrum disorders and Atypical Development. *Developmental Neurohabilitation, 14*, 221-229.

Hodgetts, S., Nicholas, D., Zwaigenbaum, L. (2013a). Home sweet home: families' experience with aggression in autism. *Focus on Autism and Other Developmental Disabilities, 28* (3), 166–174.

Horner, R. H., & Albin, R. W. (1988). Research on general-case procedures for learners with severe disabilities. *Education and Treatment of Children, 11*(4), 375–388.

Horner, R. H., & Carr, E. G. (1997). Behavioral support for students with severe disabilities functional assessment and comprehensive intervention. *The Journal of Special Education, 31*(1), 84-104.

Horner, R. H., Day, H. M., & Day, J. R. (1997). Using neutralizing routines to reduce problem behaviors. Journal of Applied Behavior Analysis, 30 (4), 601–614.

Horner R. H., Carr E., Strain P. S., Todd A. W., Reed H. K. (2002). Problem Behavior Interventions of Young Children with Autism: A research Synthesis. *The Journal of Autism and Developmental Disorders, 32,* 423-446.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional children*, 71(2), 165-179.

Hsieh, H.-H., Wilder, D. A., & Abellon, O. E. (2011). The effects of training on caregiver implementation of incidental teaching. *Journal of Applied Behavior Analysis*, *44*(1), 199–203.

Iwata B. A., Pace G. M., Dorsey M. F., Zarcone J. R., Vollmer T. R., Smith R. G., Rodgers T. A., Lerman D. C., Shore B. A., Mazaleski J. L., Goh H. L., Cowdery G. E., Kalsher M. J., McCosh K. C., Willis K. D. (1994). The Functions of Self-Injurious Behaviour: An Experimental-Epidemiological Analysis. *Journal of Applied Behavior Analysis*, *27*(2), 215-240.

Iwata, B., & DeLeon, I. (2005). The functional analysis screening tool. Gainesville, FL: The Florida Center on Self-Injury, University of Florida.

Iwata, B. A., DeLeon, I. G., & Roscoe, E. M. (2013). Reliability and validity of the functional analysis screening tool. *Journal of Applied Behavior Analysis*, 46(1), 271-284.

Kahn, S., Ingvarsson, E. T., Quigg, A. M., Seckinger, K. E., Teichman, H. M. (2011). Defining and Measuring Behaviour IN Fisher, W. W., Piazza, C. C., & Roane, H. S. Handbook of applied behavior analysis. New York: Guilford Press.

Kanne, S. M., Mazurek, M.O. (2011). Aggression in children and adolescents with ASD: Prevalence and risk factors. *Journal of Autism and developmental Disorders, 41* (7), 926-937.

Kelly, M. B. (1977). A review of the observational data-collection and reliability procedures reported in the Journal of Applied Behavior Analysis. *Journal of Applied Behavior Analysis, 10(1),* 97-101.

Kelley, M. E., Berg, W. K., Harding, J. W., Cooper-Brown, L. J. (2011) Indirect Behavioural Assessments: Interviews and Rating Scales IN Fisher, W. W., Piazza, C. C., & Roane, H. S. Handbook of applied behavior analysis. New York: Guilford Press.

Kennedy, C. H., Meyer, K. A., Knowles, T., Shukla, S. (2000). Analyzing the multiple functions of stereotypical behavior for students with autism: Implications for assessment and treatment. *Journal of Applied Behavior Analysis, 33* (4), 559–571.

Koegel, L.K., Koegel, R.L., Ashbaugh, K., Bradshaw, J. (2014). The importance of early identification and intervention for children with or at risk for autism spectrum disorder. *International Journal of Speech-Language Pathology, 6* (1), 50-56.

Kovshoff, H, Hastings R, P, Remington, B. (2011). Two-year outcomes for children with autism after the cessation of early intensive behavioral intervention. *Behaviour Modification, 35* (5), 427-50.

Lang, R., Didden, R., Machalicek, W., Rispoli, M., Sigafoos, J., Lancioni, G. (2009). Behavioral treatment of chronic skin-picking in individuals with developmental disabilities: A systematic review. *Research in Developmental Disabilities, 31*, 304–315.

Lawlor, B. K., & Hornyak, M. J. (2012). Smart Goals: How the application of smart goals can contribute to achievement of student learning outcome. *Developments in Business Simulation and Experiential Learning, 39,* 260.

Lecavalier, L., Leone, S., Wiltz, J. (2006). The impact of behaviour problems on caregiver stress in young people with autism spectrum disorders. *Journal of Intellectual Disability Research, 50* (3), 172-183.

Lerman, D. C., Dittlinger, L. H., Fentress, G., & Lanagan, T. (2011). A Comparison of Methods for Collecting Data on Performance During Discrete Trial Teaching. *Behavior Analysis in Practice, 4(1),* 53–62.

Love J. R, Carr J. E., Leblanc L. A. (2009). Functional assessment of problem behaviour in children with autism spectrum disorders: a summary of 32 outpatient cases. *Journal of Autism and Developmental Disorders, 39,* 363-372.

Lucyshyn, J. M., Albin, R. W., Horner, R. H., Mann, J. C., Mann, J. A., & Wadsworth, G. (2007). Family implementation of positive behavior support for a child with Autism: Longitudinal, single-case, experimental, and descriptive replication and extension. *Journal of Positive Behavior Interventions, 9*(3), 131-150.

Meadan-Kaplansky, H., Ostrosky, M. M., Zaghlawan, H.Y., & Yu, S. Y. (2009). Promoting the social and communicative behavior of young children with Autism Spectrum Disorders: A review of parent-implemented intervention studies. *Topics in Early Childhood Special Education, 29,* 90-104.

McIntyre L. L., Blacher J., Baker B. L. (2002). Behaviour/mental health problems in young adults with intellectual disability: the impact on families. *Journal of Intellectual Disability Research, 46*, 239–249.

McLaughlin, T. W., Denney, M. K., Snyder, P. A., Welsh. J. L. (2012). Behavior support interventions implemented by families of young children: examination of contextual fit. *Journal of Positive Behaviour Interventions, 14 (*2) 87-97.

Ministry of Children and Youth Services. (2011). Guidelines: Applied Behaviour Analysis-based Services and Supports for Children and Youth with ASD. Retrieved from http://www.auditor.on.ca/en/content/annualreports/arreports/en13/301en13.pdf.

Montgomery, J., Martin, T., Shooshtari, S., Stoesz, B. M., Heinrichs, D. J., North, S., Douglas, J. (2014). Interventions for Challenging Behaviours of Students with Autism Spectrum Disorders and Developmental Disabilities: A Synthesis Paper. *Exceptionality Education International, 23*(1), 2-21.

National Autism Centre: May Institute. (2015). National Standards Project, Phase 2. Retrieved from http://www.nationalautismcenter.org/national-standards-project/results-reports/.

Olive, M. L., Franco, J. H. (2008). (Effect) size matters: And so does the calculation. The *Behavior Analyst Today*, *9* (1), 5-10.

Olive, M. L., Smith, B. W. (2005). Effect size calculations and single subject designs. *Educational Psychology: An International Journal of Experimental Educational Psychology*, *25*(2-3), 313-324.

Oliver, A. C., Pratt, L. A. and Normand, M. P. (2015). A survey of functional behavior assessment methods used by behavior analysts in practice. *Journal of Applied Behaviour Analysis*, 48, 817–829.

Ontario Ministry of Child and Youth Service (2016). http://www.children.gov.on.ca/htdocs/English/topics/specialneeds/autism/guidelines/guidelines-2011.aspx#t3.4

Parker, R. I., Brossart, D. F. (2003). Evaluating single-case research data: A comparison of seven statistical methods, *Behavior Therapy, 34* (2), 189-211.

Parker, R. I., Brossart, D. F., Vannest, K. J., & Long, J. R. (2005). Effect sizes in single case research: How large is large? *School Psychology Review, 34*(1), 116.

Parsons, M. B., Rollyson, J. H., & Reid, D. H. (2012). Evidence-Based Staff Training: A Guide for Practitioners. *Behavior Analysis in Practice, 5*(2), 2–11.

Perry, A. Cummings, A., Dunn Geier, J., Freeman, N., Hughes, S., LaRose, L., Managhan, T., Reitzel, J., & Williams, J. (2008). Effectiveness of Intensive Behavioral Intervention in a large, community-based program. *Research in Autism Spectrum Disorders*, *2*, 621-642.

Posavick, E. (2015). Program Evaluation: Methods and Case Studies. Haboken: Taylor and Francis.

Plavnick, J. B., Ferreri, S. J., & Maupin, A. N. (2010). The effects of self-monitoring on the procedural integrity of a behavioral intervention for young children with developmental disabilities. *Journal of Applied Behavior Analysis, 43*, 2, 315–320. http://doi.org/10.1901/jaba.2010.43-315

Public Health Agency of Canada. (2012). Status Report-National Epidemiologic Database for the Study of Autism in Canada (NEDSAC). Retrieved from http://www.phac-aspc.gc.ca/publicat/hpcdp-pspmc/32-2/ar-04-eng.php.

Rapp, J. T. (2007). Further evaluation of methods to identify matched stimulation. *Journal of Applied Behavior Analysis, 40*(1), 73-88.

Rollins, P. R., Campbell, M., Hoffman, R.T., & Self, K. (2015). A Translational Community-Based Early Intervention Program for Toddlers with Autism Spectrum Disorders. *Autism: The International Journal of Research and Practice*, 1-4.

Rovai, A. P. (2003). A practical framework for evaluating online distance education programs. *The Internet and Higher Education*, *6*(2), 109-124.

Sharp, W. G., Burrell T. L., Jaquess D. L. (2014). The Autism MEAL Plan: A parent-training curriculum to manage eating aversions and low intake among children with autism. *Autism, 8* (6), 712-722.

Slavin, R., Smith, D. (2009).The relationship between sample sizes and effect sizes in systematic reviews in education. *Educational Evaluation and Policy Analysis, 31* (4), 500-506.

Smith, R. G. (2011) Developing Antecedent Interventions for Problem Behaviour IN Fisher, W. W., Piazza, C. C., & Roane, H. S. Handbook of applied behavior analysis. New York: Guilford Press.

Sarokoff, R. A., & Sturmey, P. (2004). The effects of behavioral skills training on staff implementation of discrete-trial teaching. *Journal of Applied Behavior Analysis, 37*(4), 535-538.

Suess, A. N., Romani, P. W., Wacker, D. P., Dyson, S. M., Kuhle, J. L., Lee, J. F., Waldron, D. B. (2014). Evaluating the treatment fidelity of parents who conduct in-home functional communication training with coaching via telehealth. *Journal of Behavioral Education, 23*(1), 34-59.

Thompson, R. H., & Iwata, B. A. (2007). A comparison of outcomes from descriptive and functional analyses of problem behavior. *Journal of Applied Behavior Analysis, 40*(2), 333-338.

Thompson, R. H., Borrero, J. C. (2011). Direct Observation IN Fisher, W. W., Piazza, C. C., & Roane, H. S. Handbook of applied behavior analysis. New York: Guilford Press.

Virués-Ortega J. (2010). Applied behavior analytic intervention for autism in early childhood: Meta-analysis, meta-regression and dose–response meta-analysis of multiple outcomes. *Clinical Psychology Review*, *30*(4), 387–399

Vismara, L. A., Young, G. S., Rogers, S. J. (2012). Telehealth for expanding the reach of early autism training to parents. *Autism Research and Treatment*.

Wegener, D. T., & Petty, R. E. (1994). Mood management across affective states: The hedonic contingency hypothesis. *Journal of Personality and Social Psychology, 66*, 1034-1048.

Wilder, D. A., Atwell, J., & Wine, B. (2006). The Effects of Varying Levels of Treatment Integrity on Child Compliance during Treatment with a Three-Step Prompting Procedure. *Journal of Applied Behavior Analysis*, *39*(3), 369–373.

Zablotsky, B., Black, L. I., Maenner, M. J., Schieve, L. A., Blumberg, S. J. (2015). Estimated prevalence of autism and other developmental disabilities following questionnaire changes in the 2014 National Health Interview Survey. National Health Statistics Reports, 87. http://www.cdc.gov/nchs/data/nhsr/nhsr087.pdf.

Appendix A BATT

BEHAVIOURAL ASSESSMENT AND TREATMENT TAXONOMY (BATT)

Client ID:_____ D.O.B: _____ Gender:_____

Target Behaviour	Graphed? ✓ ×
Appropriate (to increase):	
Inappropriate (to decrease):	

Please review each item on the measure and indicate if it was utilized in the file you are reviewing.

ASSESSMENT CATEGORY	
(1) NECESSARY ASSESSMENTS REVIEWED:	Documented ✓ ×
Medical "FSA-cited report by Dr B in May 2012"	
Psychiatric	
Developmental	
Academic	
Cognitive	
Communication	
Occupational Therapy	
Sensory Assessment	
Other:	
Current School Status	
Current Medical Status	

(2) MEDICATION USE

Name of Medication	Purpose	Dosage	Prescribed By?	Start Date DD/MM/YYYY	Discontinued DD/MM/YY

(3) SKILLS ASSESSMENTS:	Used ✓ × n/a

Adaptive Behaviour (e.g., VINELAND II, SIB-R, BASC):	
Children Sleep Habits Questionnaire	
Communication (e.g., VB MAPP, PPVT, EVT, CELF):	
Daily Living Skills (e.g., ABLLS):	
Functional Communication Record	
Indirect preference assessment with caregiver/child	
Screening Tool for Feeding Problems	
Sexuality (e.g., Boundaries):	
Skill Specific (e.g., Toileting):	
Social Skills (e.g., SSRS):	
Social Skills Rating System (Parent/Student - not entre section)	
Toileting Training readiness Checklist	
VB-MAPP (Bathing, Grooming, Dressing - not entire section)	
Other:	
(4) BEHAVIOURAL / PSYCHIATRIC MEASURES	Used ✓ × n/a
Dual Diagnosis Screen (e.g., Reiss, ADD)	
Severity Scale of Problem Behaviour (e.g, BPI, Aberrant Beh Checklist)	
Specific Diagnostic Screen (e.g., Y-BOCS, Conner's, BDI)	
QOL	
Other:	
(5) FUNCTIONAL BEHAVIOURAL ASSESSMENTS	Used ✓ × n/a
Indirect Measures	
Interviews	
Behavioural Contextual Interview (semi-structured)	
Bio-Psycho-Social Interview (semi-structured)	
Other:	
Questionnaires	
Questions About Behavioural Function (QABF)	
FAST	
MAS	

Target Behaviour Rating Scale (FIDD)	
Other:	
Direct Measures	
ABC Incident Recording BY Consultant/Tech	
ABC Incident Recording BY Caregiver/Mediator	
ABC Incident Recording Unspecified	
Functional Assessment Observation Form	
Informal in-vivo probe (i.e., *"Behaviour consultant asked caregiver to_____"*)	
Other:	
Experimental Methods (e*vidence of written description of conditions required*)	
Functional Analogue Assessment (antecedent & consequence manipulations) - FREQUENCY	
Functional Analogue Assessment (antecedent & consequence manipulations) - LATENCY	
Structural Analogue Assessment (antecedent manipulations)	
Other:	

(6) FBA RESULTS

Specific Measure Used (e.g., QABF)	**Escape**	**Attention**	**Tangible**	**Sensory**	**Physical**	**Automatic**		**Social**	
						R+	**R-**	**R+**	**R-**
Overall Hypothesized Function: (1 = Primary, 2 = Secondary)									

TREATMENT CATEGORY

(1) ANTECEDENT STRATEGIES:

(2) TEACHING ADAPTIVE BEHAVIOURS / SKILL BUILDING:	
SKILL (e.g., requesting _____, toilet training, etc.)	**METHOD** (e.g., Chaining, Shaping, Prompt Hierarchy, etc.)

(3) CONCEQUENCE STRATEGIES:

(4) MEDIATOR TRAINING	Used ✓ × n/a
Behaviour Skill Training (BST) Model	
Instruction/Reviewed Written Program	
Modeling – Live	
Modeling - Video	
Role-play / Practice	
Feedback	
Other:	
Adherence Measure _____% Adherence after training	
DATA COLLECTION CATEGORY	

MONITORING OF TREATMENT: (Recording system used after treatment introduced)

CASE SUMMARY SECTION					
COMPLETE FOR BEHAVIOUR TARGETTED FOR INCREASE/DECREASE					
GOAL:					
DIRECTION OF DESIRED EFFECT (circle):	INCREASE		DECREASE		
TYPE OF DATA COLLECTED: (e.g. Frequency, duration including interval such as seconds, minutes etc, interval, partial interval, rating scale, percentage occurrence [include dimension i.e. frequency, duration etc.])					
TYPE OF DATA GRAPHED: (i.e. Daily, Weekly)					
BEHAVIOUR 1: _____					
UP TO 5 **BASELINE** DATA POINTS (LAST 5):					
UP TO 5 **TREATMENT** DATA POINTS (LAST 5):					
UP TO 5 **FOLLOW UP** DATA POINTS (LAST 5):					
BEHAVIOUR 2: _____ (*if applicable*)					
UP TO 5 **BASELINE** DATA POINTS (LAST 5):					
UP TO 5 **TREATMENT** DATA POINTS (LAST 5):					
UP TO 5 **FOLLOW UP** DATA POINTS (LAST 5):					